



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 124854

**TO: Manjunath N Rao
Location: REM-3B81/3C70
Art Unit: 1652
Friday, June 18, 2004**

Case Serial Number: 10/038723

**From: Mary Jane Ruhl
Location: Biotech-Chem Library
Remsen 1-A-62
Phone: 571-272-2524**

maryjane.ruhl@uspto.gov

Search Notes

Examiner Rao,

Here are the results for your recent search request.

Please feel free to contact me if you have any questions about these results.

Thank you for using STIC services. We appreciate the opportunity to serve you.

Sincerely,

Mary Jane Ruhl
Technical Information Specialist
STIC
Remsen 1-A-62
Ext. 22524

THIS PAGE LEFT BLANK

124854

From: Chan, Christina
Sent: Wednesday, June 16, 2004 5:28 PM
To: Rao, Manjunath N.; STIC-Biotech/ChemLib
Subject: RE: RUSH sequence search request for 10/038,723

Please rush. Thanks Chris

Chris Chan

TC 1600 New Hire Training Coordinator and SPE 1644 & 1642
(571)-272-0841
Remsen, 3E89

-----Original Message-----

From: Rao, Manjunath N.
Sent: Wednesday, June 16, 2004 10:51 AM
To: Chan, Christina
Subject: RUSH sequence search request for 10/038,723

Hello Christina,

Please authorize the request below as RUSH. The reason being , this is an amended case and due this bi-week.

Thanks
-Manjunath

From: Manjunath N. Rao
Art Unit 1652, Room 3B81
Mail Box in Room 3C70
Phone: 272-0939

Date: 6-16-04

Please search the following as soon as possible for application with serial number
10/038,723

1. SEQ ID NO: 2 and amino acids 1-295 against all commercial amino acid databases,
issued patents/published applications database and pending application

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: _____
Date Completed: _____
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

TYPE OF SEARCH:
NA Sequences: _____
AA Sequences: _____
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: _____
WWW/Internet: _____
Other (specify): _____

THIS PAGE LEFT BLANK

database. Please provide a **print** of all results.

If you have any questions please call me at the above phone number.

Thanks

Manjunath N. Rao
Art Unit 1652, Room 3B81,
Mail Box in Room 3C70,
Remsen Building, USPTO
400, Dulany St.
Alexandria, VA.
Phone: 571-272-0939

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: _____
Date Completed: _____
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

TYPE OF SEARCH:
NA Sequences: _____
AA Sequences: _____
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: _____
WWW/Internet: _____
Other (specify): _____

THIS PAGE LEFT BLANK

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2004, 17:43:18 ; Search time 20 Seconds

(without alignments)
2568.314 Million cell updates/sec

Title: US-10-038-723-2

Sequence: 1 MSFRLSLALSGLVCTGLANV.....SKTTATASKSTTTRSGMSL 534

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2771	100.0	534	2 A29776	glucan 1,4-alpha-g
2	2742	99.0	640	1 ALASGR	glucan 1,4-alpha-g
3	2742	99.0	640	2 A29166	glucan 1,4-alpha-g
4	2620.5	94.6	639	2 J00607	glucan 1,4-alpha-g
5	2588.5	93.4	639	2 J00479	glucan 1,4-alpha-g
6	1935	69.8	612	2 J01346	glucan 1,4-alpha-g
7	1577	56.9	626	2 S36364	glucan 1,4-alpha-g
8	1577	56.9	626	2 T49625	glucan 1,4-alpha-g
9	1497	54.0	493	2 J06538	glucan 1,4-alpha-g
10	1321	47.7	616	2 S33908	glucan 1,4-alpha-g
11	781.5	28.2	450	2 T39433	glucan 1,4-alpha-g
12	724	26.1	604	2 J09001	glucan 1,4-alpha-g
13	676.5	24.4	519	1 A5349	glucan 1,4-alpha-g
14	597	21.5	549	1 S48474	glucan 1,4-alpha-g
15	522	18.8	778	1 ALBYG	glucan 1,4-alpha-g
16	521	18.8	767	1 J00474	glucan 1,4-alpha-g
17	221.5	8.0	615	2 A64501	glucan 1,4-alpha-g
18	181	6.5	1588	2 A86036	probable adhesin Z
19	181	6.5	1588	2 H91188	probable adhesin E
20	172	6.2	622	2 G90250	glucan 1,4 alpha g
21	170.5	6.2	1275	2 T33369	hypothetical prote
22	166.5	6.0	107	2 B60754	glucan 1,4-alpha-g
23	161	5.8	1063	2 T46731	hypothetical prote
24	160	5.8	888	2 T46726	secreted acid phos
25	159	5.7	1210	2 A25547	ice nucleation pro
26	159	5.7	12271	2 F90073	hypothetical prote
27	158	5.7	1283	2 T39174	hypothetical serin
28	157	5.7	1258	2 J00188	ice nucleation pro
29	156.5	5.6	2232	2 T34434	hypothetical prote

30	155	5.6	1322	2 S07053	ice nucleation pro
31	154.5	5.6	1200	1 SNPSO	ice nucleation pro
32	153	5.5	1131	2 T41144	hypothetical serin
33	151.5	5.5	1034	2 UC2143	ice nucleation act
34	151.5	5.5	1306	2 S25370	MSB2 protein - yea
35	151.5	5.5	1331	2 A48954	mannan endo-1,4-be
36	150	5.4	1609	2 S25345	probable membrane
37	149	5.4	967	2 S66852	hypothetical prote
38	147	5.3	507	2 S64507	probable membrane
39	146	5.3	612	2 C90419	glucan 1,4 alpha g
40	144.5	5.2	614	2 F66719	hypothetical prote
41	144	5.2	948	2 T11678	hypothetical prote
42	142	5.1	1419	2 T30531	agglutinin-like ad
43	141.5	5.1	644	1 T40712	endo-1,4-beta-xy/a
44	141.5	5.1	1441	2 B86807	hypothetical prote
45	140.5	5.1	1567	2 S11672	ice nucleation pro

ALIGNMENTS

RESULT 1

A29776
glucan 1,4-alpha-glucosidase (EC 3.2.1.3) G2 precursor - Aspergillus awamori
N:Alternate names: glucamylase G2
C:Species: Aspergillus awamori
C:Date: 20-Jun-1989 #sequence_revision 20-Jun-1989 #text_change 12-Jun-2003
C:Accession: A93066; A29776
R:Numberg, J.H.; Meade, J.H.; Cole, G.; Lawyer, F.C.; McCabe, P.; Schweickart, V.; Tal,
Mol. Cell. Biol. 4, 2306-2315, 1984
A:Title: Molecular cloning and characterization of the glucamylase gene of Aspergillus
A:Reference number: A93066; MUID:8508534; PMID:6440004
A:Accession: A93066
A:Molecule type: DNA
A:Residues: 1-534 <NDN>
A:Cross-references: GB:K02465; NID:9454405; PIDN:AB59297.1; PID:9166506
R:Numberg, J.H.; Meade, J.H.; Cole, G.; Lawyer, F.C.; McCabe, P.; Schweickart, V.; Tal,
submitted to GenBank, February 1985
A:Reference number: A94514
A:Contents: annotation; revisions to the DNA sequence and coding regions for G2 form
C:Comment: The DNA sequence was obtained from GenBank, release 55.0.
C:Superfamily: Glucan 1,4-alpha-glucosidase, Yeast type; Glucan 1,4-alpha-glucosidase ho
C:Keywords: alternative splicing; extracellular protein; glycoprotein; glycosidase; hydr
F:23-447/Domain: glucan 1,4-alpha-glucosidase homology <GAG>

Query Match	Score	DB 2;	Length	534;
Best Local Similarity	100.0%;	Pred. No. 5.9e-179;		
Matches	534;	Conservative	0;	Mismatches 0; Indels 0; Gaps 0;
QY	1	MSFRLSLALSGLVCTGLANVSKRATLDSWLSNEATVARTALINNIGADGAVSGADSCI	60	
DB	1	MSFRLSLALSGLVCTGLANVSKRATLDSWLSNEATVARTALINNIGADGAVSGADSCI	60	
QY	61	VVASPTNDPDPYFTWTRDSGLVLTIVDPRNGGTSLSLTBNYISQAIVQGISNSPG	120	
DB	61	VVASPTNDPDPYFTWTRDSGLVLTIVDPRNGGTSLSLTBNYISQAIVQGISNSPG	120	
QY	121	DLSSGAGLGEKPFVDETAATYGSWGRPORDPALRATAMIGFQWLDNGYSTATDIW	180	
DB	121	DLSSGAGLGEKPFVDETAATYGSWGRPORDPALRATAMIGFQWLDNGYSTATDIW	180	
QY	181	PLVRNDLSYVAQYNNQGTQYDLMEEVNGSSFTTIAVQHRALVBSGFATAVSSGSCWCSCQ	240	
DB	181	PLVRNDLSYVAQYNNQGTQYDLMEEVNGSSFTTIAVQHRALVBSGFATAVSSGSCWCSCQ	240	
QY	241	APETICVYQSWTGSFILANPDSRSRGKANTLIGSHTPPEAACDSTQPCSPRALA	300	
DB	241	APETICVYQSWTGSFILANPDSRSRGKANTLIGSHTPPEAACDSTQPCSPRALA	300	
QY	301	NHKEVDSFRSITLNDGLSDSEAVAVGRYPEDTYNNGNPFCLTAAAEQLYDALYQWD	360	
DB	301	NHKEVDSFRSITLNDGLSDSEAVAVGRYPEDTYNNGNPFCLTAAAEQLYDALYQWD	360	

Qy 361 KQSLVETDVSLDFPKALYSDAATGYSSSSSTYSIVDAVKTFADGFVSIIVETHAASNG 420
Db 361 KQSLVETDVSLDFPKALYSDAATGYSSSSSTYSIVDAVKTFADGFVSIIVETHAASNG 420
Qy 421 SMSEQYDKSDGEOLASADLTWVSVAALLTANNRRNSVVPASMGETSASSVPGTCAATSAIG 480
Db 421 SMSEQYDKSDGEOLASADLTWVSVAALLTANNRRNSVVPASMGETSASSVPGTCAATSAIG 480
Qy 481 TYSSTVTSWPSIVATGCTTTTATPTGSGSVTSSTKTATASKTSTTTRSGMSL 534
Db 481 TYSSTVTSWPSIVATGCTTTTATPTGSGSVTSSTKTATASKTSTTTRSGMSL 534

RESULT 2

ALASGR

glucan 1,4-alpha-glucosidase (EC 3.2.1.3) precursor - Aspergillus niger
N/Alternate names: 1,4-alpha-D-glucan glucosylhydrolase; glucosylase
C/Species: Aspergillus niger
C/Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 02-Jun-2003
C/Accession: A90986; A91161; A05287; A22149; A25402
R/Boel, E.; Hansen, M.T.; Hjort, I.; Hoegh, I.; Fill, N.P.
EMBO J. 3, 1581-1585, 1984

A/Title: Two different types of intervening sequences in the glucosylase gene from Aspe
A/Reference number: A90986; MUID:84261458; PMID:6204865
A/Accession: A90986
A/Molecule type: DNA
A/Residues: 1-640 <BOE>
A/Cross-references: GB:X00712; GB:K02466; NID:g2342; PIDN:CA25103.1; PID:g2343

A/Note: the authors translated the codon GAT for residue 317 as Asn
R/Svensson, B.; Larsen, K.; Gunnarsson, A.
Eur. J. Biochem. 154, 497-502, 1986
A/Title: Characterization of a glucosylase G2 from Aspergillus niger.
A/Reference number: A91161; MUID:86136085; PMID:3081341
A/Contents: comparison of forms G1 and G2
A/Accession: A91161
A/Molecule type: protein
A/Residues: 25-640 <SVB>
C/Comment: The large molecular form G1 is shown.
C/Comment: Smaller molecular forms of the enzyme, G2, arise by proteolytic cleavage(s) c
tive towards soluble poly- and oligosaccharides.
C/Genetics: 72/1; 167/3; 200/1; 412/3

A/Intons: 72/1; 167/3; 200/1; 412/3
C/Superfamily: glucan 1,4-alpha-glucosidase with starch-binding domain; glucan 1,4-alpha
C/Keywords: alternative splicing; extracellular protein; glycoprotein; glycosidase; hyd
F:1-24/Domain: signal sequence #status predicted <SIG>
F:23-447/Domain: glucan 1,4-alpha-glucosidase homology <AG>
F:25-640/Product: glucan 1,4-alpha-glucosidase G1 #status experimental <GG1>
F:25-538/Product: glucan 1,4-alpha-glucosidase G2, short form #status experimental <GG2>
F:25-536/Product: glucan 1,4-alpha-glucosidase G2, long form #status experimental <GG3>
F:195-419/Binding site: carboxylate (Asn) (covalent) #status experimental
F:465-467,468,477,483,484,492,508,510,513,515,522,525,527,529,530,532/Binding site: cat
F:476,486,488,496,499,500,501,502,504,506,512,514,517,518,520,524,526,528,531,534,535/Bl

Query Match 99.0%; Score 2742; DB 1; Length 640;

Best Local Similarity 99.6%; Pred. No. 6.8e-177;

Matches 528; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MSFRLSLALSGVCTGLANVISKRATLDSWLSNEATVARTAILNIGADGAWVSGADSGI 60
Db 1 MSFRLSLALSGVCTGLANVISKRATLDSWLSNEATVARTAILNIGADGAWVSGADSGI 60
Qy 61 VVASPTDNDPYFTYWRDGLVLTIVLDFRNGDTSLSLTINYSIAQAIIVQGISNPSG 120
Db 61 VVASPTDNDPYFTYWRDGLVLTIVLDFRNGDTSLSLTINYSIAQAIIVQGISNPSG 120
Qy 121 DLSSGAGLGPKNVDEATYATGSGWRPQDGPALRATAMIGFGOMLLDNGYTSTADTIW 180
Db 121 DLSSGAGLGPKNVDEATYATGSGWRPQDGPALRATAMIGFGOMLLDNGYTSTADTIW 180
Qy 121 DLSSGAGLGPKNVDEATYATGSGWRPQDGPALRATAMIGFGOMLLDNGYTSTADTIW 180
Db 121 DLSSGAGLGPKNVDEATYATGSGWRPQDGPALRATAMIGFGOMLLDNGYTSTADTIW 180
Qy 181 PIVRNDLSYVAQYWNQGYDLMEEVNGSSPFTIYVQHRALVEGSAFATAVAGSSCSWCDQ 240
Db 181 PIVRNDLSYVAQYWNQGYDLMEEVNGSSPFTIYVQHRALVEGSAFATAVAGSSCSWCDQ 240

Qy 241 APEILCYLQSFMTGSLIANFDSRSGKANTLLGSIHTFDPBAACDSTFPOCSPALA 300
Db 241 APEILCYLQSFMTGSLIANFDSRSGKANTLLGSIHTFDPBAACDSTFPOCSPALA 300
Qy 301 NHEVVDSPRSITLNDGLSDSEAVAVGRYPEDTYNNGNWFPLCTLAARQYDALYQWD 360
Db 301 NHEVVDSPRSITLNDGLSDSEAVAVGRYPEDTYNNGNWFPLCTLAARQYDALYQWD 360
Qy 361 KQSLVETDVSLDFPKALYSDAATGYSSSSSTYSIVDAVKTFADGFVSIIVETHAASNG 420
Db 361 KQSLVETDVSLDFPKALYSDAATGYSSSSSTYSIVDAVKTFADGFVSIIVETHAASNG 420
Qy 421 SMSEQYDKSDGEOLASADLTWVSVAALLTANNRRNSVVPASMGETSASSVPGTCAATSAIG 480
Db 421 SMSEQYDKSDGEOLASADLTWVSVAALLTANNRRNSVVPASMGETSASSVPGTCAATSAIG 480
Qy 481 TYSSTVTSWPSIVATGCTTTTATPTGSGSVTSSTKTATASKTSTTTRSGMSL 530
Db 481 TYSSTVTSWPSIVATGCTTTTATPTGSGSVTSSTKTATASKTSTTTRSGMSL 530

RESULT 3

A29166 glucan 1,4-alpha-glucosidase (EC 3.2.1.3) precursor - Aspergillus awamori

N/Alternate names: 1,4-alpha-D-glucan glucosylhydrolase precursor; glucosylase precursor
C/Species: Aspergillus awamori
C/Date: 08-Mar-1989 #sequence_revision 08-Mar-1989 #text_change 02-Jun-2003
C/Accession: A29166

R/Numberg, J.H.; Meade, J.H.; Cole, G.; Lawyer, F.C.; McCabe, P.; Schweickart, V.; Tal, f
Mol. Cell. Biol. 4, 2306-2315, 1984
A/Title: Molecular cloning and characterization of the glucosylase gene of Aspergillus
A/Reference number: A93066; MUID:85085934; PMID:6440004
A/Accession: A29166
A/Molecule type: DNA
A/Residues: 1-640 <NUN>
A/Cross-references: GB:K02465; NID:9454405; PIDN:AAB59296.1; PID:g166505

A/Note: See also PIR:ALASGR.
C/Comment: See also PIR:ALASGR.
C/Genetics: 72/1; 167/3; 200/1; 398/3
C/Superfamily: glucan 1,4-alpha-glucosidase with starch-binding domain; glucan 1,4-alpha
C/Keywords: alternative splicing; extracellular protein; glycoprotein; glycosidase; hyd
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-640/Product: glucan 1,4-alpha-glucosidase homology <AG>
F:23-447/Domain: glucan 1,4-alpha-glucosidase homology <AG>

Query Match 99.0%; Score 2742; DB 2; Length 640;
Best Local Similarity 99.6%; Pred. No. 6.8e-177;
Matches 528; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MSFRLSLALSGVCTGLANVISKRATLDSWLSNEATVARTAILNIGADGAWVSGADSGI 60
Db 1 MSFRLSLALSGVCTGLANVISKRATLDSWLSNEATVARTAILNIGADGAWVSGADSGI 60
Qy 61 VVASPTDNDPYFTYWRDGLVLTIVLDFRNGDTSLSLTINYSIAQAIIVQGISNPSG 120
Db 61 VVASPTDNDPYFTYWRDGLVLTIVLDFRNGDTSLSLTINYSIAQAIIVQGISNPSG 120
Qy 121 DLSSGAGLGPKNVDEATYATGSGWRPQDGPALRATAMIGFGOMLLDNGYTSTADTIW 180
Db 121 DLSSGAGLGPKNVDEATYATGSGWRPQDGPALRATAMIGFGOMLLDNGYTSTADTIW 180
Qy 181 PIVRNDLSYVAQYWNQGYDLMEEVNGSSPFTIYVQHRALVEGSAFATAVAGSSCSWCDQ 240
Db 181 PIVRNDLSYVAQYWNQGYDLMEEVNGSSPFTIYVQHRALVEGSAFATAVAGSSCSWCDQ 240
Qy 241 APEILCYLQSFMTGSLIANFDSRSGKANTLLGSIHTFDPBAACDSTFPOCSPALA 300
Db 241 APEILCYLQSFMTGSLIANFDSRSGKANTLLGSIHTFDPBAACDSTFPOCSPALA 300
Qy 301 NHEVVDSPRSITLNDGLSDSEAVAVGRYPEDTYNNGNWFPLCTLAARQYDALYQWD 360
Db 301 NHEVVDSPRSITLNDGLSDSEAVAVGRYPEDTYNNGNWFPLCTLAARQYDALYQWD 360

Qy 361 KQSLVETDVLDFPKALYSDAATGTVSSSSSTVSIVDAVKTFFADGFSIVETHAASNG 420
 Db 361 KQSLVETDVLDFPKALYSDAATGTVSSSSSTVSIVDAVKTFFADGFSIVETHAASNG 420
 Qy 421 SMSSEYDKSDGQSLARLTWSYALTLNRRNSVVPASGEGTSASSVPGTCAATSAIG 480
 Db 421 SMSSEYDKSDGQSLARLTWSYALTLNRRNSVVPASGEGTSASSVPGTCAATSAIG 480
 Qy 481 TYSSVTVSWPSIVATGTTTATPTGSGSVTSSTKTATATKSTTTTS 530
 Db 481 TYSSVTVSWPSIVATGTTTATPTGSGSVTSSTKTATATKSTTTTS 530

RESULT 4
 J00607
 glucan 1,4-alpha-glucosidase (EC 3.2.1.3) precursor - Aspergillus sp.
 N:Alternate names: glucosylase
 C:Species: Aspergillus sp.
 C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 02-Jun-2003
 C:Accession: J00607
 R:Shibuya, I.; Gomi, K.; Iimura, Y.; Takahashi, K.; Tamura, G.; Hara, S.
 Agric. Biol. Chem. 54, 1905-1914, 1990
 A:Title: Molecular cloning of the glucanase gene of Aspergillus shirousami and its ex
 A:Reference number: J00607; MUID:91182400; PMID:1368603
 A:Accession: J00607
 A:Molecule type: DNA
 A:Residues: 1-639 <SHI>
 A:Experimental source: strain RIB 2504
 C:Comment: This enzyme catalyzes the release of glucose from the non-reducing ends of st
 C:Genetics:
 A:Introns: 72/1; 166/3; 199/1; 411/3
 C:Superfamily: glucan 1,4-alpha-glucosidase with starch-binding domain; glucan 1,4-alpha
 C:Keywords: glycosidase; hydrolase; polysaccharide degradation
 F:1-24/Domain: signal sequence #status predicted <SIG>
 F:23-446/Domain: signal sequence #status predicted <SIG>
 F:25-639/Product: glucoamylase #status predicted <MAT>

Query Match 94.6%; Score 2620.5; DB 2; Length 639;
 Best Local Similarity 94.3%; Pred. No. 1e-168; Indels 1; Gaps 1;
 Matches 500; Conservative 16; Mismatches 13; Indels 1; Gaps 1;

Qy 1 MSFRLALSGVCTGLANVISKRAITLDSWLSNEATVARTALINNIGADGAVSGADSGI 60
 Db 1 MSFRLALSGVCTGLANVISKRAITLDSWLSNEATVARTALINNIGADGAVSGADSGI 60
 Qy 61 VVASPTNDPFIYTWTRDSGLVLTVDLFRNGDTSLSTIENTISAQAIYQGISNPSG 120
 Db 61 VVASPTNDPFIYTWTRDSGLVLTVDLFRNGDTSLSTIENTISAQAIYQGISNPSG 120
 Qy 121 DLSGAGGEPKENVDETAITGSGWRPORDGPALATAMIGGQWLLNGYTSTATDIY 180
 Db 121 DLSGAGGEPKENVDETAITGSGWRPORDGPALATAMIGGQWLLNGYTSTATDIY 180
 Qy 121 DLSGAGGEPKENVDETAITGSGWRPORDGPALATAMIGGQWLLNGYTSTATDIY 180
 Db 121 DLSGAGGEPKENVDETAITGSGWRPORDGPALATAMIGGQWLLNGYTSTATDIY 180
 Qy 181 PLVRNDLSYVAQYNNQGTGYDLMEEVNGSSFFTIAYQHRALVEGSAFATAVSGSCWCSQ 240
 Db 181 PLVRNDLSYVAQYNNQGTGYDLMEEVNGSSFFTIAYQHRALVEGSAFATAVSGSCWCSQ 240
 Qy 240 APELICYSQFWTGSFIILANPDSRSRGKDTWLLGSHITPDEACDSTPOPCSPRALA 299
 Db 240 APELICYSQFWTGSFIILANPDSRSRGKDTWLLGSHITPDEACDSTPOPCSPRALA 299
 Qy 301 NHKEVVDLFRSITLNDGLSDSEAVAVGRYPEDTYNNPNPFLCTLAABOLYDALYQMD 360
 Db 301 NHKEVVDLFRSITLNDGLSDSEAVAVGRYPEDTYNNPNPFLCTLAABOLYDALYQMD 360
 Qy 361 KQSLVETDVLDFPKALYSDAATGTVSSSSSTVSIVDAVKTFFADGFSIVETHAASNG 420
 Db 361 KQSLVETDVLDFPKALYSDAATGTVSSSSSTVSIVDAVKTFFADGFSIVETHAASNG 420
 Qy 421 SMSSEYDKSDGQSLARLTWSYALTLNRRNSVVPASGEGTSASSVPGTCAATSAIG 480
 Db 421 SMSSEYDKSDGQSLARLTWSYALTLNRRNSVVPASGEGTSASSVPGTCAATSAIG 480
 Qy 481 TYSSVTVSWPSIVATGTTTATPTGSGSVTSSTKTATATKSTTTTS 530
 Db 481 TYSSVTVSWPSIVATGTTTATPTGSGSVTSSTKTATATKSTTTTS 530

Qy 481 TYSSVTVSWPSIVATGTTTATPTGSGSVTSSTKTATATKSTTTTS 530
 Db 481 TYSSVTVSWPSIVATGTTTATPTGSGSVTSSTKTATATKSTTTTS 530

RESULT 5
 J00479
 glucan 1,4-alpha-glucosidase (EC 3.2.1.3) precursor - Aspergillus awamori
 N:Alternate names: glucosylase I
 C:Species: Aspergillus awamori
 C:Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 02-Jun-2003
 C:Accession: J00479
 R:Hayashida, S.; Kuroda, K.; Ohta, K.; Kuhara, S.; Fukuda, K.; Sakaki, Y.
 Agric. Biol. Chem. 53, 923-929, 1989
 A:Title: Molecular cloning of the glucanase I gene of Aspergillus awamori var. kawachi
 A:Reference number: J00479
 A:Accession: J00479
 A:Molecule type: DNA
 A:Residues: 1-639 <HAY>
 A:Experimental source: var. kawachi
 C:Genetics:
 A:Gene: GAI
 A:Introns: 72/1; 166/3; 199/1; 411/3
 C:Superfamily: glucan 1,4-alpha-glucosidase with starch-binding domain; glucan 1,4-alpha
 C:Keywords: glycosidase; hydrolase; polysaccharide degradation
 F:1-24/Domain: signal sequence #status predicted <SIG>
 F:23-446/Domain: signal sequence #status predicted <SIG>
 F:25-639/Product: glucan 1,4-alpha-glucosidase #status predicted <MAT>
 F:494-536/Region: raw-starch-affinity region

Query Match 93.4%; Score 2588.5; DB 2; Length 639;
 Best Local Similarity 93.8%; Pred. No. 1.5e-166; Indels 1; Gaps 1;
 Matches 497; Conservative 15; Mismatches 17; Indels 1; Gaps 1;

Qy 1 MSFRLALSGVCTGLANVISKRAITLDSWLSNEATVARTALINNIGADGAVSGADSGI 60
 Db 1 MSFRLALSGVCTGLANVISKRAITLDSWLSNEATVARTALINNIGADGAVSGADSGI 60
 Qy 61 VVASPTNDPFIYTWTRDSGLVLTVDLFRNGDTSLSTIENTISAQAIYQGISNPSG 120
 Db 61 VVASPTNDPFIYTWTRDSGLVLTVDLFRNGDTSLSTIENTISAQAIYQGISNPSG 120
 Qy 121 DLSGAGGEPKENVDETAITGSGWRPORDGPALATAMIGGQWLLNGYTSTATDIY 180
 Db 121 DLSGAGGEPKENVDETAITGSGWRPORDGPALATAMIGGQWLLNGYTSTATDIY 180
 Qy 121 DLSGAGGEPKENVDETAITGSGWRPORDGPALATAMIGGQWLLNGYTSTATDIY 180
 Db 121 DLSGAGGEPKENVDETAITGSGWRPORDGPALATAMIGGQWLLNGYTSTATDIY 180
 Qy 181 PLVRNDLSYVAQYNNQGTGYDLMEEVNGSSFFTIAYQHRALVEGSAFATAVSGSCWCSQ 240
 Db 181 PLVRNDLSYVAQYNNQGTGYDLMEEVNGSSFFTIAYQHRALVEGSAFATAVSGSCWCSQ 240
 Qy 240 APELICYSQFWTGSFIILANPDSRSRGKDTWLLGSHITPDEACDSTPOPCSPRALA 299
 Db 240 APELICYSQFWTGSFIILANPDSRSRGKDTWLLGSHITPDEACDSTPOPCSPRALA 299
 Qy 301 NHKEVVDLFRSITLNDGLSDSEAVAVGRYPEDTYNNPNPFLCTLAABOLYDALYQMD 360
 Db 301 NHKEVVDLFRSITLNDGLSDSEAVAVGRYPEDTYNNPNPFLCTLAABOLYDALYQMD 360
 Qy 361 KQSLVETDVLDFPKALYSDAATGTVSSSSSTVSIVDAVKTFFADGFSIVETHAASNG 420
 Db 361 KQSLVETDVLDFPKALYSDAATGTVSSSSSTVSIVDAVKTFFADGFSIVETHAASNG 420
 Qy 421 SMSSEYDKSDGQSLARLTWSYALTLNRRNSVVPASGEGTSASSVPGTCAATSAIG 480
 Db 421 SMSSEYDKSDGQSLARLTWSYALTLNRRNSVVPASGEGTSASSVPGTCAATSAIG 480
 Qy 481 TYSSVTVSWPSIVATGTTTATPTGSGSVTSSTKTATATKSTTTTS 530
 Db 481 TYSSVTVSWPSIVATGTTTATPTGSGSVTSSTKTATATKSTTTTS 530

glucan 1,4-alpha-glucosidase (EC 3.2.1.3) - Aspergillus oryzae
M:Alternate names: glucosylase
C:Species: Aspergillus oryzae
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 02-Jun-2003
C:Accession: J01346
R:Hara, Y., Tsuchiya, K., Kitamoto, K., Gomi, K., Kumagai, C., Tamura, G., Hara, S.
Gene 108, 145-150, 1991
A:Title: Nucleotide sequence and expression of the glucosylase-encoding gene (glfa) from
A:Reference number: J01346; MUID:92104497; PMID:1761224
A:Accession: J01346
A:Molecule type: DNA
A:Residues: 1-612 <HAT>
A:Cross-references: GS:D10638; DDBJ:D01108; NID:g1160312; PIDN:BA01540.1; PID:g1160313
A:Note: the authors translated the codon TTT for residue 213 as Tyr
C:Comment: This enzyme hydrolyzes starch to glucose.
C:Genetics:
A:Gene: glfa
A:Introns: 75/1; 169/3; 202/1; 414/3
C:Superfamily: glucan 1,4-alpha-glucosidase with starch-binding domain; glucan 1,4-alpha-
C:Keywords: glycosidase; hydrolase; polysaccharide degradation
A:EC:3.2.1.3
A:EC:126-443/Domain: glucan 1,4-alpha-glucosidase homology <GAG>

Query Match	69.8%;	Score 1935;	DB 2;	length 612;
Best Local Similarity	70.3%;	Pred. NO. 1.4e-122;		
Matches 367;	Conservative	59;	Mismatches 90;	Indels 6;
				Gaps 4;

Qy	1	MFSPSLILASGLVCTGGLA--NVI	SKRATLDSMLINEATVATALLNNIGADGAVNSG	58		
Dd	2	VSFSSCLRALLAGSSVLA	VQVPLRQATGLDPLWMLSTEANF	SQALILNNIGADGSAQGNP 61		
Qy	59	GIIVASPESTNDPYFTY	WTTRDSGLVLTVDL	LFNNGDTSLISTENYISACAI	VQGISNP 118	
Dd	62	GVVIAISPEKSPDPDYFT	WTTRDSGLVMKTVL	DLFFGSGADLLPIIEEFISSQARI	QGISNP 121	
Qy	119	SGDISSGAGLEPPKPNV	DETAYTGSWGR	PORDGALRATAMIGHGM	LDNGYSTATDI 178	
Dd	122	SGALSSG--GLSEPKPNV	DETAFITGAKGR	PQDGDALRLRITAMISGEM	LVENSHSITATDL 180	
Qy	179	VMPVLRNDLSVYAQW	NOGYDLMEEVNGSS	FFTIAVOHRLVEGSA	FATAVGSSCGMCD 238	
Dd	181	VMPVLRNDLSVYAQW	SQSGCFDLWEEVQGT	FFITVAVSHRALV	GSSFAKVGSSCYCD 240	
Qy	239	SOAPEILICYLOS	PFMTGSPILANFDS	SRSKXANTLLGS	IHFDEEACDOSTPQPCSPRA 298	
Dd	241	SOAPEKVRCTYLOS	PFMTGSIQANFGG	SRSGKQINIVLGS	IHFDPQATCDDTTPQPCSARA 300	
Qy	299	LANHKEVVDSPRS	IYTLNLGDSDBA	VAVGRYPEDTYN	NGNPMFLCTLAAAEQYDALYQ 358	
Dd	301	LANHKVVVDSPRS	IYALINSGRAENQ	AVAGRYPEDSYN	NGNPMFLTLTAAAEQYDALYQ 360	
Qy	359	WDKQGSLEVT	DVSLDFPKALYSDA	ATGYSSSSSTYS	IYDAVYTFADPFVSIYETHAAS 418	
Dd	361	WDKIGSLAT	TVVSLPEFFALYSSA	TGYASTTVYKQI	YSAVAVADGVQIYOTAAAS 420	
Qy	419	NGMSSEQDKSDGE	OLSARDLTW	SYAALLTANNRN	NSVPAAMEETAS	SVPTCAATSA 478
Dd	421	TGSAAEQYTKTDG	SGTSARDLTW	SYAALLTANNRN	NAVPAAMEETATATSP	ASCSTTSA 480
Qy	479	IGTYSVTVIS	MPESIVATGCT--TTT	ATPFGSGSYTS	SKTT 518	
Dd	481	SGTYSVTVITSM	PTISGYPA	ADSPCOVPT--TVST	TPAVKAT 521	

RESULT 7

glucan 1,4- α -glucosidase (EC 3.2.1.3) precursor - *Neurospora crassa*
 N/Alternate names: glucosidase; glycosylase
 C/Species: *Neurospora crassa*
 C/Date: 31-Dec-1993 #sequence revision 31-Dec-1993 #text_change 02-Jun-2003
 C/Accession: S36364; S13710; S13711; S25539
 R/Stone, P.J.; Makoff, A.J.; Parish, J.H.; Radford, A.
 Curr. Genet. 24, 205-211, 1993
 A/TITLE: Cloning and sequence analysis of the glucosylase gene of *Neurospora crassa*

A:Reference number: S36364; MUID:94037144; PMID:8221928
A:Accession: S36364
A:Molecule type: DNA
A:Residues: 1-626 <STO>
A:Cross-references: EMBL:X67291
R:Koh-Laur, S.I.; Parish, J.H.; Bleasby, A.J.; Pappin, D.J.C.; Ainley, K.; Johansen, F.E
Enzyme Microb. Technol. 11, 692-695, 1989
A:title: Exported proteins of Neurospora crassa: 1-glucosylase.
A:Reference number: S13710
A:Accession: S13710
A:Molecule type: protein
A:Residues: 36-60,'X',62,'X',64-65 <KOH>
R:Koh-Laur, S.I.; Parish, J.H.; Bleasby, A.J.; Pappin, D.J.C.; Ainley, K.; Johansen, F.E
Submitted to the Protein Sequence Database, January 1990
A:Description: Exported proteins of Neurospora crassa 1: - glucosylase.
A:Reference number: S13711
A:Accession: S13711
A:Molecule type: protein
A:Residues: 36-60,'X',62,'X',64-65 <KOW>
C:Genetics:
A:Gene: gla-1
A:Introns: 82/2
C:Superfamily: glucan 1,4-alpha-glucosidase with starch-binding domain; glucan 1,4-alpha
C:Keywords: glycosidase; hydrolase; polysaccharide degradation
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-35/Domain: propeptide #status predicted <PRO>
F:33-458/Domain: glucan 1,4-alpha-glucosidase homology <GAG>
F:336-626/Product: glucan 1,4-alpha-glucosidase #status experimental <MAT>

Query Match	56.9%	Score 1577;	DB 2;	Length 626;
Best Local Similarity	57.6%	Pred. No. 1.9e-98;		
Matches 306;	Conservative 85;	Mismatches 130;	Indels 10;	Gaps 6;

```

0Y 2 $P$R$I$L$A$T$G$V$C$T$G$L$A$N$Y$S$K$R$A$T$L$D$S$W$L$S$N$E$V$T$A$R$T$A$1$H$N$N$G$D$G$A$V$S$G$D$S$IV 61
Db 13 $F$O$V$A$T$G$L$D$P$L$H$E$K$H$S$D$I$K$R$-S$V$D$S$Y$Q$T$E$T$P$1$A$Q$K$U$L$N$G$A$S$G$C$A$S$G$A$S$V 71
0Y 62 V$P$S$T$D$N$D$P$E$Y$T$W$T$R$D$S$G$L$V$K$L$V$D$L$F$R$N$G$-D$T$S$L$T$E$N$Y$1$A$O$A$V$O$G$1$S$N$S$G 120
Db 72 V$A$P$E$K$S$S$P$D$Y$T$W$T$R$D$A$L$V$T$K$L$V$D$E$F$T$N$Y$N$T$L$Q$W$T$1$Q$A$Y$A$A$Q$A$L$Q$C$V$N$S$P 131
0Y 121 D$L$S$G$A$G$E$P$K$R$N$V$D$E$T$A$Y$T$G$S$W$G$R$P$Q$R$D$G$P$A$L$R$A$T$A$M$1$G$F$Q$W$L$L$D$N$G$Y$T$A$T$D$1$W 180
Db 122 S$L$N$G$A$G$E$P$K$R$M$V$D$L$Q$F$T$G$A$M$G$R$P$Q$R$G$P$L$R$A$L$A$L$1$G$Y$K$W$L$V$N$G$A$D$T$A$K$S$1$I$W 191
0Y 181 P$L$Y$R$D$L$S$V$A$O$Y$W$N$O$T$G$Y$D$L$M$E$V$N$G$S$S$F$T$1$A$V$O$H$R$A$L$V$E$G$A$F$A$T$A$V$S$S$C$G$C$S$Q 240
Db 192 P$L$V$K$D$L$A$T$A$Y$O$W$N$N$T$G$P$D$L$M$E$B$V$N$S$S$F$T$1$A$S$H$R$A$L$V$E$G$A$F$A$K$S$V$S$C$A$C$A$P 251
0Y 241 A$P$E$1$C$Y$O$S$P$M$T$G$S$-F$1$L$A$N$F$D$S$R$S$G$K$D$A$N$T$L$G$1$H$F$D$P$A$C$D$D$S$T$F$O$C$S$P$A$L 299
Db 252 A$P$O$L$L$C$F$Q$O$S$E$F$N$S$S$Y$1$1$S$N$F$V$N$S$G$K$D$1$N$S$1$H$N$F$D$P$A$G$C$D$V$N$F$O$C$S$D$R$A$L 311
0Y 300 A$N$H$E$V$D$S$F$R$S$1$Y$T$L$N$D$G$L$D$S$E$A$V$A$V$G$R$Y$P$E$D$Y$Y$N$G$N$P$W$F$L$C$T$1$A$A$E$O$L$Y$D$A$L$Y$Q$W 359
Db 312 A$N$H$K$V$V$D$S$M$R$-F$W$G$V$N$S$G$R$T$A$G$K$A$A$V$G$R$A$D$V$Y$N$G$N$P$W$1$A$T$1$A$A$E$O$L$Y$D$A$V$Y$W 370
0Y 360 D$K$G$S$L$E$V$T$D$V$S$L$D$F$E$K$A$L$Y$S$D$A$T$G$T$Y$S$S$S$T$Y$S$1$V$D$A$V$K$T$P$A$D$G$V$S$1$V$E$T$H$A$S$N 419
Db 371 K$G$O$S$1$T$V$T$S$L$A$F$K$D$L$V$P$S$V$T$G$T$Y$S$S$S$S$T$Y$T$A$1$1$N$A$V$T$Y$A$D$G$F$V$D$1$V$A$Q$Y$T$P$D 430
0Y 420 G$M$S$E$O$Y$D$K$S$D$E$Q$L$S$A$R$D$L$T$W$S$Y$A$L$L$T$A$N$R$N$S$V$P$P$S$W$G$T$S$A$S$V$G$T$C$A$T$S$A$I 479
Db 431 G$S$L$A$Q$F$D$D$O$S$A$P$L$S$A$T$H$L$T$W$S$Y$A$F$L$S$A$A$R$R$A$G$1$V$P$S$W$G$A$S$A$N$S$L$G$S$C$S$A$T$Y$A 490
0Y 480 G$Y$S$S$W$T$V$S$W$B$S$1$V$A$T$G$T$T$T$-A$T$P$T$G$S$-----V$T$S$K$T$A$T$A$S$K 524
Db 491 G$S$Y$A$T$A$T$A$S$P$A$N$L$T$P$A$S$1$T$V$P$T$Q$G$A$H$E$V$L$V$T$N$E$K$Y$T$S$1$G$Q$T 541

```

RESULT 8

T49625
glucan 1,4-alpha-glucosidase [imported] - *Neurospora crassa*

N:Alternate names: protein B5022.70
C:Species: Neurospora crassa
C>Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2003
C:Accession: T49625
R:Schulte, U.; Aign, V.; Hohenisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura, submitted to the Protein Sequence Database, May 2000
A:Reference number: Z25022
A:Accession: T49625
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-626 <SCH>
A:Cross-references: EMBL:AL359932; GSPDB:GN00116; NCSP:B5022.70
A:Experimental source: BAC clone B5022, strain OR74A
C:Genetics:
A:Gene: NCSP:B5022.70
A:Map position: 6
A:Insertion: 82/1
C:Superfamily: glucan 1,4-alpha-glucosidase with starch-binding domain; glucan 1,4-alpha-glucanase

Query Match 56.9%; Score 1577; DB 2; Length 626;
Best Local Similarity 57.6%; Pred. No. 1.9e-98;
Matches 306; Conservative 85; Mismatches 130; Indels 10; Gaps 6;

QY 2 SFRSLALSGVCTGLANVISKRATLDSWLSNEATVARTAILNINIGADGAWYSGADSGIV 61
DB 13 ARQAVLGJPDPLHEKRHSDIIR-SVDSYIQTETPTIAQKLLCNIGASGCRASGAASGV 71

QY 62 VASPTDNDFYTTWRDSGLVLTVDLFRNG-DTSLSTIENTYISAQAIYQGISNPSG 120
DB 72 VASPSKSPDYWTYTRDALVTKLIVDEFNDYNTTLONTIOAVAAQAKQGVSNPSG 131

QY 121 DLSAGAGEPKENVDETAYTSGMRPDRDGPALRATAMIGFQWLLDNGYSTATDIY 180
DB 132 SLNSNGAGLEPRFVNDLQOFTAGWGPQRDGPRLALIGYKMLVNSGVADTKASIT 191

QY 181 PLVRNDLSVYAQYNNQGYDLMEEVNGSSFTIAYOHRALVGSAPATVAGSSCWCDQ 240
DB 192 PIVKNDLAVTAQYNNQGYDLMEEVNGSSFTIAYOHRALVGSAPATVAGSSCWCDQ 251

QY 241 APEILCIYQSFMTGS-FILANFDSRSRSGKANTLLGSIHTPPEAACDDSTFQPCSPRAL 299
DB 252 APQILCFQOOSFWSNSGYIISFNFNYSRSGKDINSVLTSIHNPDAAGCDVNTFQPCSDRAL 311

QY 300 ANHKVNSFRSITLNLGSLDSEAVAVGRYPEDTYNNPFWFLCTLAABEOLYDALY 359
DB 312 ANHKVNVDSMR-FWGNVSGRTAGKAAAVGRYADYVNNPFWFLCTLAABEOLYDALY 370

QY 360 DKQGLLEVTVGLDFFKALYSDAATGYSSSSSTYSIYDAVKTFRADGVSIYETHAASN 419
DB 371 KKQGSITTTSTSLAFKDLVPSVSTGYSSSSSTYTAIINAVTTADGVDIVAQYTTSD 430

QY 420 GMSSEQYDKSDGEOLSARDLTWSYALLLTANNRRNSVVPASWGETSASSVPGTCAATSAI 479
DB 431 GSLAQOFKDSGAPLSATHLTWSYASFLSAAARAGIYVPSWGAASANSLLPSCSASTYA 490

QY 480 GTYSSVTVTWSYATGTTT-ATPTGSGS-----VTSTKTRATASKT 524
DB 491 GSYATATATISFPANLPASTVTPPTQTGCADHEVLTFENKVTTSYQOT 541

RESULT 9
JC6538
glucan 1,4-alpha-glucosidase (EC 3.2.1.3) precursor - Aspergillus oryzae
C:Species: Aspergillus oryzae
C>Date: 05-Dec-1998 #sequence_revision 05-Dec-1998 #text_change 12-Jun-2003
C:Accession: PC6503; JC6538
R:Hata, Y.; Ishida, H.; Ichikawa, E.; Kawato, A.; Suginami, K.; Imayasu, S.
Gene 207, 127-134, 1998
A:Title: Nucleotide sequence of an alternative glucamylase-encoding gene (glab) express
A:Reference number: JC6538; MUID:98172744; PMID:9511753
A:Accession: PC6503
A:Molecule type: protein
A:Residues: 43-53; 71-82; 136-145; 180-187; 327-336; 393-404 <HAZ>

A:Experimental source: O-1013
A:Accession: JC6538
A:Molecule type: mRNA
A:Residues: 1-493 <HAZ>
A:Cross-references: DDBJ:AB007825
C:Comment: This enzyme is used in industrial processing. It releases glucose from the no
C:Genetics:
A:Insertion: 75/1; 344/2
C:Superfamily: glucan 1,4-alpha-glucosidase, yeast type; glucan 1,4-alpha-glucosidase ho
C:Keywords: glycoprotein; glycosidase; hydrolase
F:1-26/Domain: signal sequence #status predicted <SIG>
F:26-450/Domain: glucan 1,4-alpha-glucosidase homology <GAG>
F:27-493/Product: glucan 1,4-alpha-glucosidase #status experimental <MAT>
F:139,198,255,369,384,457/Binding site: carbohydrate (am) (covalent) #status predicted

Query Match 54.0%; Score 1497; DB 2; Length 493;
Best Local Similarity 56.9%; Pred. No. 3.3e-93;
Matches 277; Conservative 80; Mismatches 128; Indels 2; Gaps 2;

QY 5 SLTALSGVCTGLANVISKRATLDSWLSNEATVARTAILNINIGADGAWYSGADSGIVAS 64
DB 8 SLNATAGVAPHSFPIHRQSDLNAPFAQPTIAQGYIYNNIGADGKLVGAAGIVAS 67

QY 65 PSTDNDFYTTWRDSGLVLTVDLFRNG-DTSLSTIENTYISAQAIYQGISNPSGSLSS 124
DB 68 PSKSNPDYFTYTRDAGLMEYIEQFIGDATTLESTIQNYVDSQANQAVSNPSGSLSD 127

QY 125 GAGLGEPRFVNDYATYSGMRPDRDGPALRATAMIGFQWLLDNGYSTATDIYPLVR 184
DB 128 GSGLAEPKFYVYISQFTDSWGRPQRDGPALRASALIAVGNLSISDKQSVVXANIMPYIQ 187

QY 185 NDLSVYAQYNNQGYDLMEEVNGSSFTIAYOHRALVGSAPATVAGSSCWCDQAPBI 244
DB 188 NDLSVYQYNNQGYDLMEEVNGSSFTIAYOHRALVGSAPATVAGSSCWCDQAC-SVAQPI 246

QY 245 LCVLQSFMTGSFILANFDS-RSGKANTLLGSIHTPPEAACDDSTFQPCSPRALANK 303
DB 247 LCHLQDFMNGSVLNSLTNGRSGIDTMSLGSITHTPPPAACDDSTFQPCSPRALSNHK 306

QY 304 EYVDSFRSITLNLGSLDSEAVAVGRYPEDTYNNPFWFLCTLAABEOLYDALY 359
DB 307 LVVDSFRSIVYGINNRRGAGKAAAVGRYADYVNNPFWFLCTLAABEOLYDALY 366

QY 364 SLEFVTVGLDFFKALYSDAATGYSSSSSTYSIYDAVKTFRADGVSIYETHAASN 423
DB 367 QNVVTEISLPFFKDLSSVNTTGYSAKSSSAVESLSAVKTYADGPIVVOEYTPDGALA 426

QY 424 EYVDSDEQOLSARDLTWSYALLLTANNRRNSVVPASWGETSASSVPGTCAATSAI 483
DB 427 EYVDSDEQOLSARDLTWSYALLLTANNRRNSVVPASWGETSASSVPGTCAATSAI 486

QY 484 SVTVTWS 490
DB 487 TPTVGSW 493

RESULT 10
S33908
glucan 1,4-alpha-glucosidase (EC 3.2.1.3) P precursor - creosote fungus
N:Alternate names: glucamylase
C:Species: Amorphotheca resinae, Hormoconis resinae (creosote fungus)
C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 02-Jun-2003
C:Accession: S33908; S36577; S31423; A60754; S31422
R:Valerio, A.E.I.; Torrell, H.T.; Tuusa, T.; Aho, S.A.; Fagerstrom, B.R.; Korhola, M.P.
Curr. Genet. 24, 38-44, 1993
A:Title: Cloning and expression of Hormoconis resinae glucamylase P cDNA in Saccharomyces
A:Reference number: S33908; MUID:93365035; PMID:8358830
A:Accession: S33908
A:Molecule type: mRNA
A:Residues: 1-616 <VALI>
A:Cross-references: EMBL:X67708; NID:g2785; PIDN:CAA47945.1; PID:g2786
A:Accession: S36577
A:Molecule type: protein

A;Residues: 30-106,138-151,168-181,217-223,260-294,310-316,321-334,336-342,369-415,423-4
R;JoutelJoki, V.; Torkkeli, T.
submitted to the EMBL Data Library, August 1992
A;Reference number: S31425
A;Accession: S31425
A;Molecule type: DNA
A;Residues: 1-616 <JOU>
A;Cross-references: EMBL:X69143; NID:G2787; PIDN:CAA48243.1; PID:G2788
R;Fagerstrom, R.; Vainio, A.; Suoranta, K.; Pakula, T.; Kalkkinen, N.; Torkkeli, H.
J. Gen. Microbiol. 136, 913-920, 1990
A;Title: Comparison of two glucoamylases from *Homocidus resiniae*.
A;Reference number: A60754; MUID:90338987; PMID:2116499
A;Accession: A60754
A;Molecule type: protein
A;Residues: 30-53,72-89,145-154,217-223,260-294,423-426, 'X', 428-436 <FAG>
A;Experimental source: strain ATCC 20495
C;Genetics:
A;Gene: gamp
A;Insertions: 76/1; 177/3; 347/2
C;Superfamily: glucan 1,4-alpha-glucosidase with starch-binding domain; glucan 1,4-alpha
C;Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
F;1-29/Domains: signal sequence #status predicted <SIG>
F;27-45/Domain: glucan 1,4-alpha-glucosidase homology <GAG>
F;30-616/Product: glucan 1,4-alpha-glucosidase #status experimental <MAT>
F;99,427/Binding site: carbohydrate (Asn) (covalent) #status experimental
F;200,514,528,587/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 47.7%; Score 1321; DB 2; Length 616;
Best Local Similarity 48.6%; Pred. No. 3,2e-81;
Matches 257; Conservative 88; Mismatches 150; Indels 34; Gaps 6;

Qy 14 CTGLANVIS-----KRATLDSWLSNEATVARTALINNGAGDAMVSGADSGIVA 63
Db 8 CAGALSDNLSLAIAPELTKARDLSSFTASRAIALQCALNNIGPDGSAVPGAGGFVA 67

Qy 64 SPSTNDNDYFTYTRDGLVLTVDLFRNGDTSLSLTENYISAQAIYVQGISNPSGD-L 122
Db 68 SPSTANPDYFTYTRDGLVLTVDLFRNGDTSLSLTENYISAQAIYVQGISNPSGF 127

Qy 123 SSGAGLSEPKFNVDETAATGSGRPPORDGPAIRATAMGFGOMLLNDGTSATDVIWPL 182
Db 128 PDGVLSEPKFNVDETAATGSGRPPORDGPAIRATAMGFGOMLLNDGTSATDVIWPL 187

Qy 183 VRNDSLYVAQYWNQGYDLMEEVNGSSEFTIYAVOHRALVEGSAFATAVGSSCSMCDQAP 242
Db 188 IANDLSYVQYWNQSGDFLMEETVYASSFTIYONHRALVEGAQALHDLGYVCTGCD-QAP 246

Qy 243 EILCYLQSFMTGSLANF--DSSRSGKANTLIGSTHPEPAACDDSTFQPCSPRALA 300
Db 247 EVLCLFLOSFMNGKTVSNINNVNNGTGDGNSIIGALSTFDIDAYCDSPFLQPCGSOLA 306

Qy 301 NHKEVDSFRSIYTLNDGLDSEBAVAVGRYPEDTYNNGNPFLLCTLLAAEQLYDALYQWD 360
Db 307 NFKVLDTFRMLYTLNAGIPGEGVAVGRVIAEDVYMGSNPFYLLTTAAEFLYDAVQWK 366

Qy 361 KQGLSEVTVDSLDFFKALYSDAATGYSS--SSSTYSIVDAVYTFADGFYSIVETHAAS 418
Db 367 ARHVLTVDETSIAFFKDIPEVTVREKSGNANSPFAIDMAYAVADSYVALIEKPIPS 426

Qy 419 NGMSSEQYDKSDGEQLSARDLTWSYALTLTANNRNSVVPASKGTASSVPGICATSA 478
Db 427 NGSLSEQFNRTGTPLSAIDLTWSSAAFTIWSORRAGQPSWSGRNMLPPTTCSAST 486

Qy 479 IGTSSVTVTWSPIVATGTTTATPTGSGVSTSTKTTATASKTST 527
Db 487 PGTY-----TPATAAGAPNVTSSCOVSTITENMAT 517

RESULT 11
T39433
glucoamylase precursor - fission yeast (*Schizosaccharomyces pombe*)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 12-Jun-2003

C;Accession: T39433
R;Xiang, Z.; Aves, S.; Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Volckaert, G.
submitted to the EMBL Data Library, March 1998
A;Reference number: Z21854
A;Accession: T39433
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-450 <LYN>
A;Cross-references: EMBL:AL022305; PIDN:CAA18423.1; GSPDB:GN00067; SPDB:SPBC14C8.05C
A;Experimental source: strain 972h-; cosmid c14C8
C;Genetics:
A;Gene: SPDB:SPBC14C8.05C
A;Map position: 2
C;Superfamily: glucan 1,4-alpha-glucosidase, yeast type; glucan 1,4-alpha-glucosidase hon

Query Match 28.2%; Score 781.5; DB 2; Length 450;
Best Local Similarity 37.6%; Pred. No. 4,4e-45;
Matches 179; Conservative 66; Mismatches 166; Indels 63; Gaps 12;

Qy 7 LALSGLVCTG---LANVISKRATLDSWLSNEATVARTALINNGAGDAMVSGADSGIVA 63
Db 8 LLLGVVASLSLSPNKRSEASMDWTQKGLAMGMNLIIGDSGMHAKDINPGCTIA 67

Qy 64 SPSTNDNDYFTYTRDGLVLTVDLFRNGDTSLSLTENYISAQAIYVQGISNPSGDLS 123
Db 68 SPSTSDPDYTYQWRDGLVLTVDLFRNGDGLSEPIYKYNDEWRLQKVPNSGDFY 127

Qy 124 SGAGLSEPKFNVDETAATGSGRPPORDGPAIRATAMGFGOMLLNDGTSATDVIWPL 183
Db 128 AG-GLSEPKFNVDETSIDGMRPPORDGPAIRATAMGFGOMLLNDGTSATDVIWPL 186

Qy 184 RNDLSYVAQYWNQGYDLMEEVNGSSEFTIYAVOHRALVEGSAFATAVGSSCSMCDQAP 243
Db 187 LADLDYVANTWTEASFDLMEIEKDVHYFTLAVQKRAMQDGTAFKRG-----APD 237

Qy 244 -----ILCYLQSFMTGSLANF--DSSRSGKANTLIGSTHPEPAACDDSTFQPCSPRALA 300
Db 238 QALYORTIEPIDLKGEPFNDPGGVTKYKGRVDRSGDLCSTLLASLYSNEPD----- 291

Qy 288 DSTFQPCSPRALANHKEVDSFRSIYTLNDGLDSEBAVAVGRYPEDTYN-----GNPMF 342
Db 292 -----MHLPLFLKQETMTDRDYPVNGMKO-----ANGRPEDVYDGVSSIGNPMF 339

Qy 343 LCTLLAAEQLYDALYQWDKQSLSEVTVDSLDFFK--ALYSDAATGYSSSSSTYSIVDA 400
Db 340 ICTSSAEIILYKALAYVNDKNGELPELTENYIHFPMKFAEPD-----PYMWSVIRKN 390

Qy 401 VKTFADGFYSIVETHAASNGSSEQYDKSDGEQLSARDLTWSYALTLTANNRNSV 456
Db 391 MHTYADNFLKAVAEFPHNGSMSBQFSDRDGHOKGARDLTWSYSSSLNALYRREAI 446

RESULT 12
JP0001
glucan 1,4-alpha-glucosidase (EC 3.2.1.3) precursor - *Rhizopus oryzae*
N;Alternate names: 1,4-alpha-D-glucan glucosylhydrolase; glucoamylase
C;Species: *Rhizopus oryzae*
C;Date: 28-Dec-1987 #sequence_revision 23-Aug-1996 #text_change 16-Jun-2000
C;Accession: JP0001
R;Ashikari, T.; Nakamura, N.; Tanaka, Y.; Kuchi, N.; Shihano, Y.; Tanaka, T.; Amachi, T.
Agric. Biol. Chem. 50, 957-964, 1986
A;Title: Rhizopus raw-starch-degrading glucoamylase: its cloning and expression in yeast
A;Reference number: A90022
A;Accession: JP0001
A;Molecule type: DNA
A;Residues: 1-604 <ASH>
A;Cross-references: GB:D00049; PIDN:BA00033.1; PID:G218035
A;Experimental source: strain SAM0034
A;Note: there are two errors in the published sequence (personal communication): GCT (11)
V to the known amino acid sequence of 11 peptides from glucoamylase, including the amino-
R;Tanaka, Y.; Ashikari, T.; Nakamura, N.; Kuchi, N.; Shihano, Y.; Amachi, T.; Yoshizumi,
Agric. Biol. Chem. 50, 965-969, 1980
A;Title: Comparison of amino acid sequences of three glucoamylases and their structure-f

submitted to the EMBL Data Library, March 1991
A>Description: Nucleotide and deduced amino acid sequence of the glucoamylase

A: Molecule type: DNA

A:Residues: 1-503,507,'W',513-514,'TG',516 <YAM>
 A:Cross-references: EMBL:M16166; NID:G172592; PIDN:AAA5042.1; PID:G172593
 R:Pardo, J.M.; Ianez, E.; Zalacain, M.; Clares, M.G.; Jimenez, A.
 FEBS Lett. 239, 179-184, 1988
 A:Title: Similar short elements in the 5' regions of the STA2 and SGA genes from Sacchar
 A:Reference number: S27281; MUID:89031230; PMID:3141213
 A:Accession: S27284
 A:Molecule type: DNA
 A:Residues: 1-183,'H',185-190 <PAR>
 A:Cross-references: EMBL:X13585; NID:G4461; PIDN:CAA32071.1; PID:G4463
 A:Gene: SGD:SGA1; MIPS:YIL099W
 A:Cross-references: SGD:S0001361; MIPS:YIL099W
 A:Map position: 9L
 C:Function:
 A:Description: hydrolysis of terminal 1,4-linked alpha-D-glucose residues successively fr
 C:Superfamily: glucan 1,4-alpha-glucosidase, yeast type; glucan 1,4-alpha-glucosidase ho
 C:Keywords: glycosidase; hydrolase; polysaccharide degradation; yeast vacuole
 F:77-536/Domain: glucan 1,4-alpha-glucosidase homology <GAG>

Query Match 21.5%; Score 597; DB 1; Length 549;
 Best Local Similarity 30.4%; Pred. No. 1.6e-32;
 Matches 153; Conservative 89; Mismatches 202; Indels 60; Gaps 14;

Qy 1 MSFSLALSGVCTGLANVISKRATLDSWLSNEATVARTAILNNIGADGAWVSGADSGI 60
 Db VQLDADVLMNGTVYD-SNGAMDSALBEMLOQKQKVSIRIPENIGSAVYPS-ILPGV 113
 Qy 61 VVASPSTNDPQFYFTWRDGLVLTVDLFRNGDTSLSLTENYISAQAIIVQGISNPSG 120
 Db 114 VIASPSQTHDPYFQWIRDSALTTINSIVS--HSAGPAIEYLLQYLVNSFHLQRTNN--- 167
 Qy 121 DLSSGAG-----LGEPRNVDEATYSGMGRPQDGPALRATAMIGFGQMLDNG--- 170
 Db 168 TLGAGIGYNTDVALGDPKMNVDNTAFTEBGRQNDGPALRSIALIKIIDIYIKQSGTDL 227
 Qy 171 -----YSTATDIWPLVRNDLSVAQYMNQGYDLMEEVNGSSFTTAVGHRALVEGSA 225
 Db 228 GAKYFPQSTA-DIFDDIVRMDLRFLIDHMNSSGEFLMEVNGMFFTLVQLSAVDSLS 286
 Qy 226 PATAVSGSCWCD--SQAPRILCYL---QSFMTGSFILANF-----DSRSKSDA 270
 Db 287 YFNASERSSPFEELRQTRDISKFLVDPANGFNGKX--NYIVGTPMADTLRSGLDI 343
 Qy 271 NTLLGSIHTPDPEAACDSTFQPCSPRALANHKEVVDSPFSIYTLNDGLDSEAVAVGRY 330
 Db 344 STLAAVTVDHAPSA-SHLPFDINDPVLNTLHLLMHMRSIYPIINDSSKNATGIALGRY 402
 Qy 331 PEDTY-----YNGNPMFLCTLAABQLYDALYQMDKQSLSEVTVDSIDFFKALYSDATG 385
 Db 403 PEDYVDGYGFGEGNPMVLATCTASTTYLYQLYRHISQHDLVVPMNNDCSNAFMSSELVFS 462
 Qy 386 TYSS-----SSSTYSSIVDAVKTRPADGVSVIVETHAASNGSMNQYDKSDGEQ 433
 Db 463 NLTLTGNDGYLLLEFNTPAFNQTIQKIFQLADSEFLVYLKAHVGTDGELSDQFNKKTGFM 522
 Qy 434 LSARDLFWSYAALLTANRRNSVY 457
 Db 523 QGAQHLTMSYTSFMDAYQIROEVL 546

RESULT 15

ALBYG

glucan 1,4-alpha-glucosidase (EC 3.2.1.3) precursor - yeast (Saccharomyces cerevisiae)
 N:Alternate names: 1,4-alpha-D-glucan glucohydrolase; glucamylase
 C:Species: Saccharomyces cerevisiae
 C>Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 05-Sep-1997
 C:Accession: A21896; A23470

R:Yamaashita, I.; Suzuki, K.; Fukui, S.
 J. Bacteriol. 161, 567-573, 1985

A:Title: Nucleotide sequence of the extracellular glucamylase gene STA1 in the yeast Sa
 A:Reference number: A21896; MUID:85104778; PMID:3918017
 A:Note: S. diastaticus

A:Accession: A21896
 A:Molecule type: DNA
 A:Residues: 1-778 <YAL>
 A:Experimental source: strain 5106-9A; ATCC 60709
 R:Yamaashita, I.; Suzuki, K.; Sakuro, F.
 Agric. Biol. Chem. 50, 475-482, 1986
 A:Title: Proteolytic processing of glucamylase in the yeast Saccharomyces cerevisiae.
 A:Reference number: A23470
 A:Contents: signal sequence cleavage site
 A:Accession: A23470
 A:Molecule type: protein
 A:Residues: 1-65 <YAZ>
 C:Genetics:
 A:Gene: STA1
 C:Superfamily: yeast glucan 1,4-alpha-glucosidase STA1; glucan 1,4-alpha-glucosidase hom
 C:Keywords: extracellular protein; glycoprotein; glycosidase; hydrolase; polysaccharide c
 F:1-32/Domain: signal sequence #status experimental <SIG>
 F:33-778/Product: glucan 1,4-alpha-glucosidase #status predicted <MPT>
 F:345-778/Domain: glucan 1,4-alpha-glucosidase homology <GAG>
 F:46,319,333,425,434,445,524,557,656,661,731,752/Binding site: carbohydrate (asn) (coval
 F:635/Active site: Asp #status predicted

Query Match 18.8%; Score 522; DB 1; Length 778;
 Best Local Similarity 30.5%; Pred. No. 2.9e-27;
 Matches 139; Conservative 80; Mismatches 177; Indels 60; Gaps 15;

Qy 1 MSFSLALSGVCTGLANVISKRATLDSWLSNEATVARTAILNNIGADGAWVSGADSGI 60
 Db 324 VQLADVLMNGTVYD-SNGAMDSALBEMLOQKQKVSIRIPENIGSAVYPS-ILPGV 381
 Qy 61 VVASPSTNDPQFYFTWRDGLVLTVDLFRNGDTSLSLTENYISAQAIIVQGISNPSG 120
 Db 382 VIASPSQTHDPYFQWIRDSALTTINSIVS--HSADPA-IEYLLQYLVNSFHLQRTNN--- 435
 Qy 121 DLSSGAG-----LGEPRNVDEATYSGMGRPQDGPALRATAMIGFGQMLDNG--- 170
 Db 436 TLGAGIGYNTDVALGDPKMNVDNTAFTEBGRQNDGPALRSIALIKIIDIYIKQSGTDL 495
 Qy 171 -----YSTATDIWPLVRNDLSVAQYMNQGYDLMEEVNGSSFTTAVGHRALVEGSA 225
 Db 496 GAKYFPQSTA-DIFDDIVRMDLRFLIDHMNSSGEFLMEVNGMFFTLVQLSAVDSLS 554
 Qy 226 PATAVSGSCWCD--SQAPRILCYL---QSFMTGSFILANF-----DSRSKSDA 270
 Db 555 YFNASERSSPFEELRQTRDISKFLVDPANGFNGKX--NYIVGTPMADTLRSGLDI 611
 Qy 271 NTLLGSIHTPDPEAACDSTFQPCSPRALANHKEVVDSPFSIYTLNDGLDSEAVAVGRY 330
 Db 612 STLAAVTVDHAPSA-SHLPFDINDPVLNTLHLLMHMRSIYPIINDSSKNATGIALGRY 670
 Qy 331 PEDTY-----YNGNPMFLCTLAABQLYDALYQMDKQSLSEVTVDSIDFFKALYSDATG 385
 Db 671 PEDYVDGYGFGEGNPMVLATCTASTTYLYQLYRHISQHDLVVPMNNDCSNAFMSSELVFS 730
 Qy 386 TYSS-----SSSTYSSIVDAVKTRPADGV 409
 Db 731 NLTLTGNDGYLLLEFNTPAFNQTIQKIFQLADSEFL 766

Search completed: June 17, 2004, 17:48:00
 Job time : 22 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2004, 17:35:48 ; Search time 59 Seconds
(without alignments)
2557.294 Million cell updates/sec

Title: US-10-038-723-2
Perfect score: 2771
Sequence: 1 MSFSLSLAGLVCTGLANV.....SKTATASKSTTTTSGMSL 534

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A.Geneseq.29Jan04:*

1: geneseqp1980a:*
2: geneseqp1990a:*
3: geneseqp2000a:*
4: geneseqp2001a:*
5: geneseqp2002a:*
6: geneseqp2003a:*
7: geneseqp2003bs:*
8: geneseqp2004a:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2771	100.0	534	2	AAV18090 Truncated
2	2771	100.0	534	2	AAV23338 Aspergill
3	2771	100.0	534	3	AAV23337 Aspergill
4	2771	100.0	534	3	AAV23337 Aspergill
5	2771	100.0	534	4	AAV23337 Aspergill
6	2771	100.0	534	4	AAV23337 Aspergill
7	2742	99.0	640	4	AAV23337 Aspergill
8	2742	99.0	640	4	AAV23337 Aspergill
9	2683.5	96.8	639	1	AAV23337 Aspergill
10	2658	95.9	630	1	AAV23337 Aspergill
11	2628	94.8	616	2	AAV23337 Aspergill
12	2625	94.7	616	2	AAV23337 Aspergill
13	2625	94.7	616	2	AAV23337 Aspergill
14	2615	94.4	616	2	AAV23337 Aspergill
15	2615	94.4	616	2	AAV23337 Aspergill
16	2614	94.3	616	3	AAV23337 Aspergill
17	2612	94.3	616	3	AAV23337 Aspergill
18	2612	94.3	616	3	AAV23337 Aspergill
19	2604	94.0	616	2	AAV23337 Aspergill
20	2604	94.0	616	2	AAV23337 Aspergill
21	2600.5	93.8	621	2	AAV23337 Aspergill
22	2589.5	93.5	621	3	AAV23337 Aspergill
23	2584.5	93.3	621	3	AAV23337 Aspergill
24	2523	89.1	1095	6	ABP96630 Alpha-amy
25	2468	89.1	471	4	ABM00045 AMG SEQ I

26	1850.5	66.8	631	6	ABB80181 A. fumiga
27	1770	63.9	624	4	AAV23339 Thermocac
28	1589	61.0	618	2	AAV23337 Talaromy
29	1674.5	60.4	591	2	AAV23337 Talaromy
30	1615.5	58.3	630	5	AAV23337 Talaromy
31	1577	56.9	626	2	AAV23337 Talaromy
32	1425.5	51.4	581	3	AAV23337 Talaromy
33	1425.5	51.4	581	3	AAV23337 Talaromy
34	1425.5	51.4	581	3	AAV23337 Talaromy
35	1408.5	50.8	704	6	ABB80170 A. fumiga
36	1380.5	49.8	620	6	ABB80170 A. fumiga
37	1321	47.7	616	6	AAV23337 Talaromy
38	725	26.2	579	6	AAV23337 Talaromy
39	724	26.1	604	1	AAV23337 Talaromy
40	641.5	22.2	624	2	AAV23337 Talaromy
41	615.5	22.2	497	1	AAV23337 Talaromy
42	600.5	21.7	918	1	AAV23337 Talaromy
43	374	13.5	293	1	AAV23337 Talaromy
44	181	6.5	1588	7	ADC01413 Enterohae
45	172	6.2	32	2	AAV23337 Talaromy

ALIGNMENTS

RESULT 1	AAV18090	standard; protein; 534 AA.
XX	AAV18090;	
AC	AAV18090;	
XX	AAV18090;	
DT	06-AUG-1999	(first entry)
XX	06-AUG-1999	(first entry)
DE	Truncated A. niger glucoamylase GI protein sequence.	
XX	Glucoamylase; saccharification; liquefied starch solution; cyclodextrin;	
KW	enzymatic saccharification; monosaccharide production; oligosaccharide;	
KW	dextrose; trehalose; isomaltoligosaccharide; maltooligosaccharide.	
XX	Aspergillus niger.	
OS	Aspergillus niger.	
PN	WO9927124-A1.	
XX	03-JUN-1999.	
PD	03-JUN-1999.	
XX	23-NOV-1998;	98WO-US024871.
PF	23-NOV-1998;	98WO-US024871.
XX	26-NOV-1997;	97US-00979673.
PR	30-JUN-1998;	98US-00107657.
XX	30-JUN-1998;	98US-00107657.
PA	(NOVO) NOVO-NORDISK AS.	
XX	(STAL) STALEY MFG CO A E.	
PI	Liaw GC, Pedersen S, Hendriksen HV, Svendsen A, Nielsen BR;	
PI	Nielsen RI.	
XX	WPI; 1999-357849/30.	
DR	N-PSDB; AA76983.	
XX	Saccharification of liquefied starch solution for production of, e.g.	
PT	monosaccharides.	
XX	Claim 18; Page 55-58; 63pp; English.	
PS	This sequence is a truncated A. niger glucoamylase GI protein. The	
XX	invention relates to a method of saccharifying a liquefied starch	
CC	solution comprising: (a) a saccharification step during which one or more	
CC	enzymatic saccharification stages (using a mutated Aspergillus niger	
CC	glucoamylase) take place; (b) one or more high temperature membrane	
CC	separation steps; and (c) re-circulation of the saccharification enzyme;	
CC	in which the membrane separation steps are carried out as an integral	
CC	part of the saccharification step. The method can be used for the	
CC	production of mono and/or oligosaccharides from starch, including	

CC dextrose, trehalose, isomaltoligosaccharides, cyclodextrins or
CC maltoligosaccharides. The use of the membrane separation step improves
CC the efficiency of the process and improves yields and purity
XX
XX Sequence 534 AA:

Query Match 100.0%; Score 2771; DB 2; Length 534;
Best Local Similarity 100.0%; Pred. No. 1.5e-214;
Matches 534; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSFSLIALSGVCTGLANVSKRATLDSWLSNEATVARTAILNINIGADGAWVSGADSGI 60
DB 1 MSFSLIALSGVCTGLANVSKRATLDSWLSNEATVARTAILNINIGADGAWVSGADSGI 60
QY 61 VVASPTDNDPYFTWTRDGLVLTVDLFRNGDTSLSSTIENYISAQAIIVQGISNPSG 120
DB 61 VVASPTDNDPYFTWTRDGLVLTVDLFRNGDTSLSSTIENYISAQAIIVQGISNPSG 120
QY 121 DLSGAGLGEPRKFNVDATAVYGSWGRPQRDGPALRATAMIGFGQWLLDNGYSTATDIW 180
DB 121 DLSGAGLGEPRKFNVDATAVYGSWGRPQRDGPALRATAMIGFGQWLLDNGYSTATDIW 180
QY 181 PLVNDLSYVAQYNNQGYDLMEEVNGSSFETIAVQHRALVEGSAFATAVSSSCSWCDSQ 240
DB 181 PLVNDLSYVAQYNNQGYDLMEEVNGSSFETIAVQHRALVEGSAFATAVSSSCSWCDSQ 240
QY 241 APEILCYLQSFMTGSFILANFDSRSRSGKANTLLGSIHTFDPEACDSTFQPCSPRALA 300
DB 241 APEILCYLQSFMTGSFILANFDSRSRSGKANTLLGSIHTFDPEACDSTFQPCSPRALA 300
QY 301 NHKEVDSFRSIYTLNDGLSDSEAVAVGRYPEDTYNNGNPMFLCTLAQAOLYDALYQMD 360
DB 301 NHKEVDSFRSIYTLNDGLSDSEAVAVGRYPEDTYNNGNPMFLCTLAQAOLYDALYQMD 360
QY 361 KQGLSEVTVDSLDFPKALYSDAATGYSSSSSTYSIVDAKTFADGVSIIVETHAASNG 420
DB 361 KQGLSEVTVDSLDFPKALYSDAATGYSSSSSTYSIVDAKTFADGVSIIVETHAASNG 420
QY 421 SMSQYKSDGEQLSARDLTWSYAALLTANNRRNSVVPASWGETSASSVPGTCAATSAIG 480
DB 421 SMSQYKSDGEQLSARDLTWSYAALLTANNRRNSVVPASWGETSASSVPGTCAATSAIG 480
QY 481 TYSSTVTSWPSIYATGTTTATPTGSGSVTSSTKTTATASKTSTTRGMSL 534
DB 481 TYSSTVTSWPSIYATGTTTATPTGSGSVTSSTKTTATASKTSTTRGMSL 534

RESULT 2
AA2338
ID AA2338 standard; protein; 534 AA.

XX AC AA2338;
XX DT 02-SEP-1999 (first entry)
XX DE Aspergillus niger glucoamylase enzyme.
XX KM Glucoamylase enzyme; saccharification; starch hydrolyase;
XX KW dextrose syrup; syrup; ethanol; fuel; beverage; citric acid;
XX KW ascorbic acid; lysine; glutamic acid.
XX OS Aspergillus niger.
XX PN MO9928448-A1.
XX PD 10-JUN-1999.
XX PF 26-NOV-1998; 98WO-DK000520.
XX PR 26-NOV-1997; 97US-00979673.
XX PR 30-DEC-1997; 97DK-00001557.
XX PR 30-JUN-1998; 98US-00107657.
XX PR 10-JUL-1998; 98DK-00000925.

XX PA (NOVO) NOVO-NORDISK AS.
XX PI Nielsen BR, Nielsen RI, Lehmbeck U;
XX DR WP1; 1999-404822/34.
XX DR N-PSDB; AAX81797.
XX PT New glucoamylase obtained from Talaromyces emersonii.
XX PS Disclosure; Page 63-64; 79pp; English.

XX The present sequence represents a glucoamylase enzyme obtained from
XX Aspergillus niger. The specification describes a Talaromyces emersonii
XX glucoamylase enzyme. The glucoamylase enzymes have high thermal stability
XX so that a saccharification process may be carried out within a shorter
XX period of time or the process may be carried out using a lower enzyme
XX dosage. The glucoamylase enzymes can be used for saccharifying starch
XX hydrolyzate for converting starch or partially hydrolysed starch into a
XX syrup containing dextrose. They can be used for producing
XX oligosaccharides, specialty syrups, ethanol for fuel, beverages or
XX organic compounds such as citric acid, ascorbic acid, lysine or glutamic
XX acid

XX Sequence 534 AA:
QY Query Match 100.0%; Score 2771; DB 2; Length 534;
DB Best Local Similarity 100.0%; Pred. No. 1.5e-214;
Matches 534; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSFSLIALSGVCTGLANVSKRATLDSWLSNEATVARTAILNINIGADGAWVSGADSGI 60
DB 1 MSFSLIALSGVCTGLANVSKRATLDSWLSNEATVARTAILNINIGADGAWVSGADSGI 60
QY 61 VVASPTDNDPYFTWTRDGLVLTVDLFRNGDTSLSSTIENYISAQAIIVQGISNPSG 120
DB 61 VVASPTDNDPYFTWTRDGLVLTVDLFRNGDTSLSSTIENYISAQAIIVQGISNPSG 120
QY 121 DLSGAGLGEPRKFNVDATAVYGSWGRPQRDGPALRATAMIGFGQWLLDNGYSTATDIW 180
DB 121 DLSGAGLGEPRKFNVDATAVYGSWGRPQRDGPALRATAMIGFGQWLLDNGYSTATDIW 180
QY 181 PLVNDLSYVAQYNNQGYDLMEEVNGSSFETIAVQHRALVEGSAFATAVSSSCSWCDSQ 240
DB 181 PLVNDLSYVAQYNNQGYDLMEEVNGSSFETIAVQHRALVEGSAFATAVSSSCSWCDSQ 240
QY 241 APEILCYLQSFMTGSFILANFDSRSRSGKANTLLGSIHTFDPEACDSTFQPCSPRALA 300
DB 241 APEILCYLQSFMTGSFILANFDSRSRSGKANTLLGSIHTFDPEACDSTFQPCSPRALA 300
QY 301 NHKEVDSFRSIYTLNDGLSDSEAVAVGRYPEDTYNNGNPMFLCTLAQAOLYDALYQMD 360
DB 301 NHKEVDSFRSIYTLNDGLSDSEAVAVGRYPEDTYNNGNPMFLCTLAQAOLYDALYQMD 360
QY 361 KQGLSEVTVDSLDFPKALYSDAATGYSSSSSTYSIVDAKTFADGVSIIVETHAASNG 420
DB 361 KQGLSEVTVDSLDFPKALYSDAATGYSSSSSTYSIVDAKTFADGVSIIVETHAASNG 420
QY 421 SMSQYKSDGEQLSARDLTWSYAALLTANNRRNSVVPASWGETSASSVPGTCAATSAIG 480
DB 421 SMSQYKSDGEQLSARDLTWSYAALLTANNRRNSVVPASWGETSASSVPGTCAATSAIG 480
QY 481 TYSSTVTSWPSIYATGTTTATPTGSGSVTSSTKTTATASKTSTTRGMSL 534
DB 481 TYSSTVTSWPSIYATGTTTATPTGSGSVTSSTKTTATASKTSTTRGMSL 534

RESULT 3
AAB03450
ID AAB03450 standard; protein; 534 AA.

XX AC AAB03450;
XX

DT 03-JAN-2001 (first entry)
 XX Aspergillus niger G2 glucoamylase.
 XX Glucoamylase G2; starch hydrolysis; high fructose corn syrup;
 KM chemical stability.
 XX Aspergillus niger.
 OS
 XX Key Location/Qualifiers
 FT Peptide 1..27
 FT /label= prepropeptide_sequence
 FT Cleavage-site 27..28
 FT /label= tripeptidyl_aminopeptidase_cleavage_site
 FT Protein 28..534
 FT /label= mature_glucoamylase
 XX
 XX WO200034452-A1.
 XX
 XX 15-JUN-2000.
 XX
 XX 07-DEC-1999; 99WO-DK00066.
 XX
 XX 07-DEC-1998; 98DK-00001616.
 XX 24-MAR-1999; 99DK-00000409.
 XX
 XX (NOVO) NOVO-NORDISK AS.
 XX
 XX Nielsen BR, Svendsen A, Boisen K, Vind J, Pedersen H;
 PI WPI; 2000-431296/37.
 XX
 XX Variants of parent fungal glucoamylase useful for producing ethanol,
 PT beverages and in fermentation processes comprise a peptide extension at
 PT the amino terminal.
 XX
 XX Disclosure; Page 50-51; 61pp; English.
 XX
 XX The present sequence is the G2 glucoamylase from Aspergillus niger. It
 CC was used to construct a version of the enzyme which has improved thermal
 CC stability, and which can be used more efficiently in starch hydrolysis.
 CC This process is used during high fructose corn syrup production, as well
 CC as in the production of ethanol for fuel or beverage, and in fermentation
 CC processes for producing organic compounds such as citric acid, ascorbic
 CC acid, lysine and glutamic acid
 CC
 XX Sequence 534 AA;
 SQ
 Query Match 100.0%; Score 2771; DB 3; Length 534;
 Best Local Similarity 100.0%; Pred. No. 1.5e-214;
 Matches 534; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 301 NHKEVVDSPRSIYTLNDGLSDSEAVAVGRYPEDTYNNGNPFLLCTLAALAEQYDALYQWD 360
 QY 361 KQGSLEVTDVSLDPFKALYSDAATGTYSSSSSTYSIDAVKTFPDPGFSIYETAASNG 420
 DB 361 KQGSLEVTDVSLDPFKALYSDAATGTYSSSSSTYSIDAVKTFPDPGFSIYETAASNG 420
 QY 421 SMSSEYDSDGSEQLSARDLUTWSYALLTANNRRNSVVPASWGETSASSVPGCAATSAG 480
 DB 421 SMSSEYDSDGSEQLSARDLUTWSYALLTANNRRNSVVPASWGETSASSVPGCAATSAG 480
 QY 481 TYSSVTWTSWPSIYATGTTTATPTPGSGSVTSSTKTATASKTSTTRSGMSL 534
 DB 481 TYSSVTWTSWPSIYATGTTTATPTPGSGSVTSSTKTATASKTSTTRSGMSL 534
 RESULT 4
 AA77740
 ID AA77740 standard; protein; 534 AA.
 XX
 XX AA77740;
 AC
 XX 22-MAY-2000 (first entry)
 DT
 XX A. niger G2 glucoamylase.
 DE
 XX Glucoamylase; variant; starch conversion; saccharification; ethanol;
 KM fuel; beverage; fermentation; citric acid; ascorbic acid; thermostable;
 KM G2 glucoamylase; fungal.
 XX
 XX Aspergillus niger.
 OS
 XX WO200004136-A1.
 XX
 XX 27-JAN-2000.
 PD
 XX 09-JUL-1999; 99WO-DK000392.
 PF
 XX 15-JUL-1998; 98DK-00000937.
 PR 17-DEC-1998; 98DK-00001667.
 XX
 XX (NOVO) NOVO-NORDISK AS.
 XX
 XX Nielsen BR, Svendsen A, Pedersen H, Vind J, Hendriksen HV;
 PI Prandsen TP;
 PI WPI; 2000-182412/16.
 DR N-PSDB; AA287842.
 DR
 XX Variant fungal glucoamylases with improved thermostability and increased
 PT specific activity, useful in saccharification processes.
 PT
 XX Claim 1; Page 78-81; 116pp; English.
 PS
 XX The invention relates to variant fungal glucoamylases comprising specific
 CC mutations. The variant comprises one or more mutation in position/region
 CC 1-18, 19-35, 40-62, 73-80, 93-127, 170-184, 200-212, 234-246, 287-319,
 CC 334-341, 353-374, 388-414, 445-470 of the parent G2 glucoamylase sequence
 CC (AA77740) from A. niger, with the exception of: N20C, A27C, S30P, Y48W,
 CC W50H, W52P, R54K/L, D55G/V, G57A, K108R, D112Y, Y116A/W, S119C/W/E/G/Y/P,
 CC W120H/L/F/Y, G121T/A, K122Y, K108R, D112Y, Y116A/W, S119C/W/E/G/Y/P,
 CC K123G, G174C, Y175F, D176N/E, L177H/D, W178R/D, E179Q/D, E180D/O,
 CC V181D/A/T, N182A/D/Q/Y/S, G183K, S184H, W212F, R241K, A246C, D253E/Q,
 CC A303V, R305K, Y306P, D309N/E, Y312W, W317F, E389D/Q, H391W, A392D, A393P,
 CC N395Q, G396S, E400Q/C, Q401E, G407D, E408P, L410F, S411A/G/C/H/D, and
 CC S460P. The glucoamylase variants are useful in a starch conversion
 CC process, especially continuous processes which include a continuous
 CC saccharification process. The variants can be used for producing
 CC oligosaccharides, especially syrups, or ethanol for fuel or beverages.
 CC They can also be used in fermentation processes for producing organic
 CC compounds such as citric acid, ascorbic acid, lysine and glutamic acid.
 CC The glucoamylase variants have improved thermostability and/or increased
 CC specific activity. This is advantageous in industrial saccharification

CC processes. The risk of microbial contamination is also reduced when
 CC carrying the saccharification process at temperatures above 63 plusOC. An
 CC increased specific activity towards short chain saccharides such as
 CC maltose (without reducing the activity towards oligosaccharides) would
 CC also permit using a lower enzyme dosage and/or shorter process times. The
 CC present sequence represents the G2 glucamylase from *A. niger*, the parent
 CC glucamylase used for constructing the variants

XX Sequence 534 AA;

Query Match 100.0%; Score 2771; DB 3; Length 534;

Best Local Similarity 100.0%; Pred. No. 1.5e-214; Mismatches 0; Gaps 0;

Matches 534; Conservative 0; Indels 0; Gaps 0;

QY 1 MSFRLSLALSGLVCTGLANVSKRATLDSWLSNEATYARATLANNIGADGAWGADSGI 60
 DB 1 MSFRLSLALSGLVCTGLANVSKRATLDSWLSNEATYARATLANNIGADGAWGADSGI 60
 QY 61 VVASPSTDNPDYFYTWTRDSGLVLTVDLFRNGDTSLSLTENYISAQAIVOGINSNPSG 120
 DB 61 VVASPSTDNPDYFYTWTRDSGLVLTVDLFRNGDTSLSLTENYISAQAIVOGINSNPSG 120
 QY 121 DLSGAGLGEPKENVDETAATYATGSGRQPDGPALRATAMIGFGQWLLDNGYTSATDIW 180
 DB 121 DLSGAGLGEPKENVDETAATYATGSGRQPDGPALRATAMIGFGQWLLDNGYTSATDIW 180
 QY 181 PLYVNDLSYVAQYWNQGYDLMEEVNGSSFFTIAYQHRALVEGSAFATAVSSCSWCDQ 240
 DB 181 PLYVNDLSYVAQYWNQGYDLMEEVNGSSFFTIAYQHRALVEGSAFATAVSSCSWCDQ 240
 QY 241 APEILCYLQSFMTGSFTILANFDSRSRSGKANTLLGSIHTDPPEAACDSTFQPCSPALA 300
 DB 241 APEILCYLQSFMTGSFTILANFDSRSRSGKANTLLGSIHTDPPEAACDSTFQPCSPALA 300
 QY 301 NHKEVDSFRSIYTLNDGLSDSEAVAGRPEDTYNGNPMFLCTLLAAEQLYDALYQWD 360
 DB 301 NHKEVDSFRSIYTLNDGLSDSEAVAGRPEDTYNGNPMFLCTLLAAEQLYDALYQWD 360
 QY 361 KQSLLEVTDVSLDFPKALYSDAATGTYSSSSSTYSIVDAKTPADGFSIVETHAASNG 420
 DB 361 KQSLLEVTDVSLDFPKALYSDAATGTYSSSSSTYSIVDAKTPADGFSIVETHAASNG 420
 QY 421 SMSFOYKSDGEQLSARDLTWVSVAALLTANNRRNSVVPASWGETSASVPGTCAATSAIG 480
 DB 421 SMSFOYKSDGEQLSARDLTWVSVAALLTANNRRNSVVPASWGETSASVPGTCAATSAIG 480
 QY 481 TYSSVTVTSWPSIYATGTTTATPTGSGVSTSTSKTTATASKTSTTRSGMSL 534
 DB 481 TYSSVTVTSWPSIYATGTTTATPTGSGVSTSTSKTTATASKTSTTRSGMSL 534

RESULT 5

AA848171 standard; procein; 534 AA.

AA848171;

02-APR-2001 (first entry)

A. niger G1 glucamylase polypeptide.

Thermoascus crustaceus; glucamylase; starch conversion; ethanol;
 maltose syrup; beverage; citric acid; ascorbic acid; detergent;
 thermostability; glucose; G1 glucamylase.

Aspergillus niger.

Location/Qualifiers

1..24

/note= "signal peptide"

25..534

/note= "mature protein"

PN W0200075296-A1.
 XX 14-DEC-2000.
 XX 02-JUN-2000; 2000WO-DK000301.
 XX 02-JUN-1999; 99DK-00000779.
 XX (NOVO) NOVO NORDISK AS.
 XX
 PI Nielsen BR, Kauppinen MS, Nielsen RI;
 XX WPI; 2001-071066/08.
 DR N-PSDB; AAC84444.
 XX
 PT Novel Thermoascus crustaceus glucamylase polypeptide useful in a
 PT continuous starch conversion process, detergents and for producing
 PT oligosaccharides, specially syrups, ethanol for fuel or drinking,
 PT beverages.
 XX
 PS Disclosure; Page 52-55; 61pp; English.

CC The invention provides a Thermoascus crustaceus glucamylase polypeptide.
 CC The glucamylase polypeptide can be expressed by standard recombinant
 CC methodology and is useful in a continuous starch conversion process, for
 CC producing oligosaccharides, specially syrups such as maltose syrups,
 CC ethanol for fuel or drinking ethanol, beverages, and organic compounds
 CC such as citric acid, ascorbic acid, lysine or glutamic acid. It is also
 CC useful in detergents such as laundry detergent compositions, dish wash
 CC compositions and/or hard surface cleaning compositions. The T. crustaceus
 CC glucamylase has higher thermostability than Aspergillus niger G1
 CC glucamylase. It also has higher specific activity and/or decreased
 CC glucose reversion tendency. The present sequence represents the A. niger
 CC G1 glucamylase

XX Sequence 534 AA;

Query Match 100.0%; Score 2771; DB 4; Length 534;

Best Local Similarity 100.0%; Pred. No. 1.5e-214; Mismatches 0; Gaps 0;

Matches 534; Conservative 0; Indels 0; Gaps 0;

QY 1 MSFRLSLALSGLVCTGLANVSKRATLDSWLSNEATYARATLANNIGADGAWGADSGI 60
 DB 1 MSFRLSLALSGLVCTGLANVSKRATLDSWLSNEATYARATLANNIGADGAWGADSGI 60
 QY 61 VVASPSTDNPDYFYTWTRDSGLVLTVDLFRNGDTSLSLTENYISAQAIVOGINSNPSG 120
 DB 61 VVASPSTDNPDYFYTWTRDSGLVLTVDLFRNGDTSLSLTENYISAQAIVOGINSNPSG 120
 QY 121 DLSGAGLGEPKENVDETAATYATGSGRQPDGPALRATAMIGFGQWLLDNGYTSATDIW 180
 DB 121 DLSGAGLGEPKENVDETAATYATGSGRQPDGPALRATAMIGFGQWLLDNGYTSATDIW 180
 QY 181 PLYVNDLSYVAQYWNQGYDLMEEVNGSSFFTIAYQHRALVEGSAFATAVSSCSWCDQ 240
 DB 181 PLYVNDLSYVAQYWNQGYDLMEEVNGSSFFTIAYQHRALVEGSAFATAVSSCSWCDQ 240
 QY 241 APEILCYLQSFMTGSFTILANFDSRSRSGKANTLLGSIHTDPPEAACDSTFQPCSPALA 300
 DB 241 APEILCYLQSFMTGSFTILANFDSRSRSGKANTLLGSIHTDPPEAACDSTFQPCSPALA 300
 QY 301 NHKEVDSFRSIYTLNDGLSDSEAVAGRPEDTYNGNPMFLCTLLAAEQLYDALYQWD 360
 DB 301 NHKEVDSFRSIYTLNDGLSDSEAVAGRPEDTYNGNPMFLCTLLAAEQLYDALYQWD 360
 QY 361 KQSLLEVTDVSLDFPKALYSDAATGTYSSSSSTYSIVDAKTPADGFSIVETHAASNG 420
 DB 361 KQSLLEVTDVSLDFPKALYSDAATGTYSSSSSTYSIVDAKTPADGFSIVETHAASNG 420
 QY 421 SMSFOYKSDGEQLSARDLTWVSVAALLTANNRRNSVVPASWGETSASVPGTCAATSAIG 480
 DB 421 SMSFOYKSDGEQLSARDLTWVSVAALLTANNRRNSVVPASWGETSASVPGTCAATSAIG 480

QY 481 TYSSTVTSMPISVATGTTTATPTGSGSVTSTSKTTATASKTTTTRSGMSL 534
DB 481 TYSSTVTSMPISVATGTTTATPTGSGSVTSTSKTTATASKTTTTRSGMSL 534

RESULT 6
AAB61904
ID AAB61904 standard; protein; 534 AA.
XX AAB61904;
AC AAB61904;
DT 08-MAY-2001 (first entry)
XX
DE A. niger G2 glucoamylase.
XX
KM Glucoamylase; variant; starch; dextrose; saccharification; maltodextrin;
KW glucose syrup; fuel; ethanol; beverage; fermentation; citric acid;
KM ascorbic acid.
XX
OS Aspergillus niger.
XX
FH Key Location/Qualifiers
FT Peptide 1..24
FT Protein /note= "signal peptide"
FT Protein 25..534
FT Protein /note= "mature protein"
FN WO200104273-A2.
XX
PD 18-JAN-2001.
XX
PF 07-JUL-2000; 2000WO-DK000373.
XX
PR 09-JUL-1999; 99DK-00000999.
XX
PA (NOVO) NOVO NORDISK AS.
PI Nielsen BR, Svendsen A, Pedersen H, Vind J, Hendriksen HV,
PI Frandsen TP;
XX
DR WPI; 2001-138334/14.
DR N-PSDB; AAC85099.
XX
PT Novel variant of parent glucoamylase useful in starch conversion process,
PT and for producing oligosaccharides, maltodextrins, glucose syrups, fuel,
PT drinking ethanol, beverage and organic compounds.
XX
PS Claim 1; Page 51-53; 58pp; English.
XX
CC The invention relates to a variant of a parent glucoamylase, comprising
CC an alteration at positions 59, 66, 72, 119, 189, 223, 227, 313, 340, 342,
CC 352, 379, 386, 393, 395, 402, 408, 416, 425, 427, 444, 486, 490, 494,
CC where the alteration is independently an insertion, substitution or
CC deletion of amino acid which occupies the position. The variant
CC glucoamylase is useful for converting starch or partially hydrolyzed
CC starch into a syrup containing dextrose, by saccharifying starch
CC hydrolyzate. The variant is useful in the starch conversion process, for
CC producing oligosaccharides, maltodextrins or glucose syrups, fuel,
CC drinking ethanol, beverage and in a fermentation process for producing
CC organic compounds, such as citric acid, ascorbic acid, lysine and
CC glutamic acid. It is useful for improving the thermal stability and/or
CC specific activity of a parent glucoamylase. The present sequence
CC represents an A. niger G2 glucoamylase, the parent enzyme from which the
CC variant of the invention is derived
XX
SQ Sequence 534 AA;

Query Match 100.0%; Score 2771; DB 4; Length 534;
Best Local Similarity 100.0%; Pred No. 1.5e-214;
Matches 534; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSFSLALSGVCTGLANVISKRATLDMSLSEATVARTAILNNIGADGAWGADSGI 60
|||||

DB 1 MSFSLALSGVCTGLANVISKRATLDMSLSEATVARTAILNNIGADGAWGADSGI 60
QY 61 VVASTSTNDPQFYMTWTDGSLVLTIVDLPENGGTSLSTENTISAOAYOGISNPSG 120
DB 61 VVASTSTNDPQFYMTWTDGSLVLTIVDLPENGGTSLSTENTISAOAYOGISNPSG 120
QY 121 DLSGAGLGEPEFNDETAATYTSWGRPORDEGALBATAIMIGFQWLDNGYSTATDIYW 180
DB 121 DLSGAGLGEPEFNDETAATYTSWGRPORDEGALBATAIMIGFQWLDNGYSTATDIYW 180
QY 181 PLVRNDLSYVAQYMNQGYDLMEBVNGSSFPTIAVQHRALVGSAPATVAGSSCSWCDSQ 240
DB 181 PLVRNDLSYVAQYMNQGYDLMEBVNGSSFPTIAVQHRALVGSAPATVAGSSCSWCDSQ 240
QY 241 APELICYSQFWTGSFILANPSSSSGDNANTLLGSIHFDEAACDOSTPQCSPRALA 300
DB 241 APELICYSQFWTGSFILANPSSSSGDNANTLLGSIHFDEAACDOSTPQCSPRALA 300
QY 301 NHKEVVDSPRSIYTLNDGLSDSEAVAVGRYPEDTYNGNPMFLCTIAAEQLYDALYQMD 360
DB 301 NHKEVVDSPRSIYTLNDGLSDSEAVAVGRYPEDTYNGNPMFLCTIAAEQLYDALYQMD 360
QY 361 KQGSLEVTDVSLDFPKALYSDAATGYSSSSSTYSIVDAVKTADGFPVSIYETHAASNG 420
DB 361 KQGSLEVTDVSLDFPKALYSDAATGYSSSSSTYSIVDAVKTADGFPVSIYETHAASNG 420
QY 421 SMSQYDYSDEQSLASBDLTWSYALLTNRRNRNVVPASWETSASSVPGCAATSAIG 480
DB 421 SMSQYDYSDEQSLASBDLTWSYALLTNRRNRNVVPASWETSASSVPGCAATSAIG 480
QY 481 TYSSTVTSMPISVATGTTTATPTGSGSVTSTSKTTATASKTTTTRSGMSL 534
DB 481 TYSSTVTSMPISVATGTTTATPTGSGSVTSTSKTTATASKTTTTRSGMSL 534

RESULT 7
AA77741
ID AA77741 standard; protein; 640 AA.
XX
AC AA77741;
XX
DT 22-MAY-2000 (first entry)
XX
DE A. niger G1 glucoamylase.
XX
KM Glucoamylase; variant; starch conversion; saccharification; ethanol;
KW fuel; beverage; fermentation; citric acid; ascorbic acid; thermostable;
XX G1 glucoamylase; fungal.
XX
OS Aspergillus niger.
XX
PN WO200004136-A1.
XX
PD 27-JAN-2000.
XX
PF 09-JUL-1999; 99WO-DK000392.
XX
PR 15-JUL-1998; 98DK-00000937.
PR 17-DEC-1998; 98DK-00001667.
XX
PA (NOVO) NOVO-NORDISK AS.
XX
PI Nielsen BR, Svendsen A, Pedersen H, Vind J, Hendriksen HV,
PI Frandsen TP;
XX
DR WPI; 2000-182412/16.
DR N-PSDB; AA87843.
XX
PT Variant fungal glucoamylases with improved thermostability and increased
PT specific activity, useful in saccharification processes.
XX
PS Disclosure; Page 91-93; 116pp; English.
XX

CC The invention relates to variant fungal glucoamylases. The variants
CC comprise specific mutations in the parent G2 glucoamylase (GmG) sequence
CC (AAV77740) from *A. niger* (see AA287842 for specific positions of the
CC mutations). The glucoamylase variants are useful in a starch conversion
CC process, especially continuous process which include a continuous
CC saccharification process. The variants can be used for producing
CC oligosaccharides, especially syrups, or ethanol for fuel or beverages.
CC They can also be used in fermentation processes for producing organic
CC compounds such as citric acid, ascorbic acid, lysine and glutamic acid.
CC The glucoamylase variants have improved thermostability and/or increased
CC specific activity. This is advantageous in industrial saccharification
CC processes. The risk of microbial contamination is also reduced when
CC carrying the saccharification process at temperatures above 63 plusOC. An
CC increased specific activity towards short chain saccharides such as
CC maltose (without reducing the activity towards oligosaccharides) would
CC also permit using a lower enzyme dosage and/or shorter process times. The
CC present sequence represents the G1 glucoamylase from *A. niger*
XX
SQ Sequence 640 AA;
Query Match 99.0%; Score 2742; DB 3; Length 640;
Best Local Similarity 99.6%; Pred. No. 4.4e-212;
Matches 528; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 MSFSLALSGLVCTGLANVISKRATLDSWLSNEATVARTAILNNGADGAWSGADSGI 60
DB 1 MSFSLALSGLVCTGLANVISKRATLDSWLSNEATVARTAILNNGADGAWSGADSGI 60
QY 61 VVASPTDNDPXYFTWTRDSGLVLTVDLFRNGDTSLSTIENYISAQAIVQGISNPSG 120
DB 61 VVASPTDNDPXYFTWTRDSGLVLTVDLFRNGDTSLSTIENYISAQAIVQGISNPSG 120
QY 121 DLSGAGLGEKPFVNDETATYTGSGWRPQRDPALRATAMIFGQWLLDNGTSTATDIW 180
DB 121 DLSGAGLGEKPFVNDETATYTGSGWRPQRDPALRATAMIFGQWLLDNGTSTATDIW 180
QY 181 PLVNDLSYVAQYNNQGYDLMEEVNGSPFTIAVQHRALVEGSAFATAVSSCSWCDSD 240
DB 181 PLVNDLSYVAQYNNQGYDLMEEVNGSPFTIAVQHRALVEGSAFATAVSSCSWCDSD 240
QY 241 APEILCYLQSFMTGSFILLANFDSRSRSGKANTLLGSIHTFDPBAACDSTFQPCSPRALA 300
DB 241 APEILCYLQSFMTGSFILLANFDSRSRSGKANTLLGSIHTFDPBAACDSTFQPCSPRALA 300
QY 301 NHKEVDSFRSITYTLNDGLDSEAVAVGRYPEDTYNNGNPMFLCTIAAEOLYDALYQMD 360
DB 301 NHKEVDSFRSITYTLNDGLDSEAVAVGRYPEDTYNNGNPMFLCTIAAEOLYDALYQMD 360
QY 361 KQSLLEVTVDSLDFFKALYSDAATGTYSSTYSIVDAVKTPADGFSIVETHAASNG 420
DB 361 KQSLLEVTVDSLDFFKALYSDAATGTYSSTYSIVDAVKTPADGFSIVETHAASNG 420
QY 421 SMSQYDSDGEOLASRDLTWSYALLTANNRRNSVVPASGERTSASSVPCTCAATSAIG 480
DB 421 SMSQYDSDGEOLASRDLTWSYALLTANNRRNSVVPASGERTSASSVPCTCAATSAIG 480
QY 481 TYSSTVTSWPSIVATGTTTATPTGSGSVTSTSKTATASKTSTTSS 530
DB 481 TYSSTVTSWPSIVATGTTTATPTGSGSVTSTSKTATASKTSTTSS 530
RESULT 8
AAB61905
ID AAB61905 standard; protein; 640 AA.
XX
AC AAB61905;
XX
DT 08-MAY-2001 (first entry)
XX
DE A. niger protein sequence id No. 13.
XX
KW Glucoamylase; variant; starch; dextrose; saccharification; maltodextrin;
glucose syrup; fuel; ethanol; beverage; fermentation; citric acid;

KW ascorbic acid.
XX
XX Aspergillus niger.
OS
XX WO200104273-A2.
XX
XX 18-JAN-2001.
PD
XX 07-JUL-2000; 2000WO-DK000373.
XX
XX 09-JUL-1999; 99DK-00000999.
PR
XX (NOVO) NOVO NORDISK AS.
PA
XX Nielsen BR, Svendsen A, Pedersen H, Vind J, Hendriksen HV;
PI Frandsen TP;
PI
XX WPI; 2001-138334/14.
DR
XX
XX Novel variant of parent glucoamylase useful in starch conversion process,
PT and for producing oligosaccharides, maltodextrins, glucose syrups, fuel,
PT drinking ethanol, beverage and organic compounds.
PS Disclosure; Page 57-58; 58pp; English.
XX
XX The invention relates to a variant of a parent glucoamylase, comprising
CC an alteration at positions 59, 66, 72, 119, 189, 223, 227, 313, 340, 342,
CC 352, 379, 386, 393, 395, 402, 408, 416, 425, 427, 444, 486, 490, 494,
CC where the alteration is independently an insertion, substitution or
CC deletion of amino acid which occupies the position. The variant
CC glucoamylase is useful for converting starch or partially hydrolyzed
CC starch into a syrup containing dextrose, by saccharifying starch
CC hydrolyzate. The variant is useful in the starch conversion process, for
CC producing oligosaccharides, maltodextrins or glucose syrups, fuel,
CC drinking ethanol, beverage and in a fermentation process for producing
CC organic compounds, such as citric acid, ascorbic acid, lysine and
CC glutamic acid. It is useful for improving the thermal stability and/or
CC specific activity of a parent glucoamylase
XX
SQ Sequence 640 AA;
Query Match 99.0%; Score 2742; DB 4; Length 640;
Best Local Similarity 99.6%; Pred. No. 4.4e-212;
Matches 528; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 MSFSLALSGLVCTGLANVISKRATLDSWLSNEATVARTAILNNGADGAWSGADSGI 60
DB 1 MSFSLALSGLVCTGLANVISKRATLDSWLSNEATVARTAILNNGADGAWSGADSGI 60
QY 61 VVASPTDNDPXYFTWTRDSGLVLTVDLFRNGDTSLSTIENYISAQAIVQGISNPSG 120
DB 61 VVASPTDNDPXYFTWTRDSGLVLTVDLFRNGDTSLSTIENYISAQAIVQGISNPSG 120
QY 121 DLSGAGLGEKPFVNDETATYTGSGWRPQRDPALRATAMIFGQWLLDNGTSTATDIW 180
DB 121 DLSGAGLGEKPFVNDETATYTGSGWRPQRDPALRATAMIFGQWLLDNGTSTATDIW 180
QY 181 PLVNDLSYVAQYNNQGYDLMEEVNGSPFTIAVQHRALVEGSAFATAVSSCSWCDSD 240
DB 181 PLVNDLSYVAQYNNQGYDLMEEVNGSPFTIAVQHRALVEGSAFATAVSSCSWCDSD 240
QY 241 APEILCYLQSFMTGSFILLANFDSRSRSGKANTLLGSIHTFDPBAACDSTFQPCSPRALA 300
DB 241 APEILCYLQSFMTGSFILLANFDSRSRSGKANTLLGSIHTFDPBAACDSTFQPCSPRALA 300
QY 301 NHKEVDSFRSITYTLNDGLDSEAVAVGRYPEDTYNNGNPMFLCTIAAEOLYDALYQMD 360
DB 301 NHKEVDSFRSITYTLNDGLDSEAVAVGRYPEDTYNNGNPMFLCTIAAEOLYDALYQMD 360
QY 361 KQSLLEVTVDSLDFFKALYSDAATGTYSSTYSIVDAVKTPADGFSIVETHAASNG 420
DB 361 KQSLLEVTVDSLDFFKALYSDAATGTYSSTYSIVDAVKTPADGFSIVETHAASNG 420

QY 421 SMSEQYDKSDGQSLARDLTWVSYALLTANNRRNSVVPASWGETSASVPGTCAATSAIG 480
 DB 421 SMSEQYDKSDGQSLARDLTWVSYALLTANNRRNSVVPASWGETSASVPGTCAATSAIG 480
 QY 481 TYSSTVTWSPSIVATGCTTTTATPTGSGSVTSTSKTTATASKSTTTTSS 530
 DB 481 TYSSTVTWSPSIVATGCTTTTATPTGSGSVTSTSKTTATASKSTTTTSS 530

RESULT 9

AAp40212
 ID AAp40212 standard; protein; 639 AA.

AC AAp40212;
 DT 25-MAR-2003 (revised)
 DT 09-JAN-1992 (first entry)
 DE Sequence encoded by A.awamori glucoamylase genomic region.
 KM Starch hydrolysis; glucose.

OS Aspergillus awamori.

PN WO8402921-A.

PD 02-AUG-1984.

PF 26-JAN-1984; 84WO-US000122.

PR 28-JAN-1983; 83US-00461920.

PR 20-DEC-1983; 83US-00563941.

PR 27-APR-1987; 87US-00564076.

PR 27-APR-1987; 87US-00047552.

PA (CETU) CETUS CORP.

PA (CETU) CETUS CORP.

PI Nunberg JH, Flatsgaard JE, Innis MA, Gelfand DH, Meade JH;

DR WPI; 1984-201413/32.

DR N-PSDB; AAN40166.

PT DNA sequence coding for fungal glucoamylase protein - for expression in

PT yeast etc. for prodn. of the enzyme.

PS Example; Table 1, Page 21-25; 66pp; English.

XX The inventors claim a modified DNA sequence coding for fungal

CC glucoamylase protein or its single or multiple base substitutions

CC deletions, insertions or inversions is new (see AAN40165). It is derived

CC from natural, synthetic or semisynthetic sources and is capable, when

CC correctly combined with a cleaved expression vector, of expressing a non-

CC native protein having glucoamylase activity on transformation of a host

CC organism by the vector. (Updated on 25-MAR-2003 to correct PI field.)

XX SQ Sequence 639 AA;

QY Query Match 96.8%; Score 2683.5; DB 1; Length 639;
 DB Best Local Similarity 98.1%; Pred. No. 2.3e-207;
 DB Matches 520; Conservative 1; Mismatches 8; Indels 1; Gaps 1;

QY 181 PLVRNDLSVAAQYMQTGYDLMEEYNGSSFTTIAVOHRAHYGSAFATAVGSSCSWCDGQ 240
 DB 180 PLVRNDLSVAAQYMQTGYDLMEEYNGSSFTTIAVOHRAHYGSAFATAVGSSCSWCDGQ 239
 QY 241 APEILCYLQSFMTGSPFIANPDSRSRGDANTLGSHTFDEPAACDDSTFOCPSPRALA 300
 DB 240 APEILCYLQSFMTGSPFIANPDSRSRGDANTLGSHTFDEPAACDDSTFOCPSPRALA 299
 QY 301 NKEEYVDSFRSITLTNDGLSDSEAVAVGRYPEDTYNGNPMFLCTLAABQLYDALYQMD 360
 DB 300 NKEEYVDSFRSITLTNDGLSDSEAVAVGRYPEDTYNGNPMFLCTLAABQLYDALYQMD 359
 QY 361 KOGSLEVTDSLDFPKALYSDAATGTSSTSSSTSIYDAVTFPAGFVSIVETAAASNG 420
 DB 360 KOGSLEVTDSLDFPKALYSDAATGTSSTSSSTSIYDAVTFPAGFVSIVETAAASNG 419
 QY 421 SMSEQYDKSDGQSLARDLTWVSYALLTANNRRNSVVPASWGETSASVPGTCAATSAIG 480
 DB 420 SMSEQYDKSDGQSLARDLTWVSYALLTANNRRNSVVPASWGETSASVPGTCAATSAIG 479
 QY 481 TYSSTVTWSPSIVATGCTTTTATPTGSGSVTSTSKTTATASKSTTTTSS 530
 DB 480 TYSSTVTWSPSIVATGCTTTTATPTGSGSVTSTSKTTATASKSTTTTSS 529

RESULT 10

AAp81876
 ID AAp81876 standard; protein; 630 AA.

AC AAp81876;
 DT 24-OCT-2003 (revised)
 DT 25-MAR-2003 (revised)
 DT 15-MAR-1992 (first entry)

DE Sequence of amylglucosidase.

XX Enzyme; brewing; bread-making; dextrin.

XX Aspergillus niger; NCI 22343.

XX EP260160-A.

XX 16-MAR-1986.

XX 10-JUN-1987; 87EP-00401300.

XX 10-JUN-1986; 86FR-00008387.

XX 13-APR-1987; 87FR-00005207.

XX 13-APR-1987; 87FR-00005208.

XX (TRGE) TRANSGENE SA.

XX Labat N, Loison G, Lemoine Y;

XX WPI; 1988-072593/11.

XX N-PSDB; AAN82019.

XX New DNA block for expressing amylglucosidase in yeast - contg. gene plus

XX transcription and export sequences, and transformed cells useful in

XX brewing and bread making.

XX Example; Fig 2; 38pp; French.

XX EcoRI-sal I fragments of Aspergillus niger NCI 22343 were cloned in

XX pBR322 and two clones contg. portions of the gene isolated using a pool

XX of three probes-TG282, TG283, TG284- (AAN82014-6) designed on the basis

XX of already published S08. The clones were designated pTG1830 and 1831. A

XX cDNA bank of A.niger NCI 22343 was screened for amylglucosidase using

XX probe TG433 (AAN82017) which corresp. to AAs 163-172 of the mature

XX protein (see AAN82019). Oligo TG387 (AAN82018) was used to confirm the

XX identity of the clones detected. (Updated on 25-MAR-2003 to correct PR

CC field.) (Updated on 25-MAR-2003 to correct PA field.) (Updated on 24-OCT-2003 to standardise OS field)

XX Sequence 630 AA;

Query Match 95.9%; Score 2658; DB 1; Length 630;
Best Local Similarity 97.2%; Pred. No. 2,6e-205;
Matches 515; Conservative 2; Mismatches 3; Indels 10; Gaps 1;

```

QY 1 MSFSLALSGVCTGLANVSKRATLDSWLSNEATVARTAILNNIGADGAWVGADSGI 60
DB 1 MSFSLALSGVCTGLANVSKRATLDSWLSNEATVARTAILNNIGADGAWVGADSGI 60
QY 61 VVASPTDNDPYFTWTRDSGLVLTIVDLFRNGDTSILSTIENYISAQAIVQGISNPSG 120
DB 61 VVASPTDNDPYFTWTRDSGLVLTIVDLFRNGDTSILSTIENYISAQAIVQGISNPSG 120
QY 121 DLSGAGLGEPEKFNVDERTAYTGSWGRPQRDPALRATMIGFGQMLDNGTSTRTDVIW 180
DB 121 DLSGAGLGEPEKFNVDERTAYTGSWGRPQRDPALRATMIGFGQMLDNGTSTRTDVIW 180
QY 181 PLVNDLSYVAQYNNQGYDLMEEVNGSSFTIAVOHRALVEGSAFATAVGSSGSCWCDSD 240
DB 171 PLVNDLSYVAQYNNQGYDLMEEVNGSSFTIAVOHRALVEGSAFATAVGSSGSCWCDSD 230
QY 241 APEILCYQSFWTGSFILANFDSRSRQKANTLGSIHFPPEACDSTFQPCSPRALA 300
DB 231 APEILCYQSFWTGSFILANFDSRSRQKANTLGSIHFPPEACDSTFQPCSPRALA 290
QY 301 NHRKVDNFRSITYLNDGLSSEAVAVGRYPEDTYNGNPMWFLCTLAABQLYDALYQMD 360
DB 291 NHRKVDNFRSITYLNDGLSSEAVAVGRYPEDTYNGNPMWFLCTLAABQLYDALYQMD 350
QY 361 KQGLSEVTVLDFPFKALYSDATGTYSSSSSTYSIVDAVKTADGVSIVETHAASNG 420
DB 351 KQGLSEVTVLDFPFKALYSDATGTYSSSSSTYSIVDAVKTADGVSIVETHAASNG 410
QY 421 SMSQYDKSDGEQLSARDLTWSYALLTANNRRNSVVPASWGETSASSVPCTCATSAIG 480
DB 411 SMSQYDKSDGEQLSARDLTWSYALLTANNRRNSVVPASWGETSASSVPCTCATSAIG 470
QY 481 TYSSVTVTWSRISYATGTTTATPTGSGTSTSKTTATASKTSTTSS 530
DB 471 TYSSVTVTWSRISYATGTTTATPTGSGTSTSKTTATASKTSTTSS 520

```

RESULT 11

AAB15176 standard; procein; 616 AA.

AC AAB15176;

DT 11-DEC-2000 (first entry)

DE Aspergillus awamori glucosylase.

KM Glucosylase; enzyme; carbohydrolase; glucose;

KW 1,4-alpha-D-glucan glucosylase.

OS Aspergillus awamori.

PN MO200043504-A1.

PD 27-JUL-2000.

PF 10-JAN-2000; 2000MO-US000532.

PR 22-JAN-1999; 99US-002336063.

PA (IOWA) UNIV IOWA STATE RES POUND INC.

PI Allen MJ, Fang T, Li Y, Liu H, Chen H, Coutinho P, Honzacko R; Ford C;

XX WPI; 2000-514725/46.

XX Fungal glucosylase for selective production of glucose rather than alpha-1,6 linked disaccharide isomaltose, has mutation pair Asn20Cys coupled with Ala27Cys forming disulfide bond between the two stabilizing members.

PS Disclosure; Page 152-153; 160pp; English.

XX The present sequence is glucosylase (1,4-alpha-D-glucan glucosylase; E.C. 3.2.1.3). This enzyme is a carbohydrolase, and cleaves D-glucose from the nonreducing ends of maltotrioligosaccharides, attacking alpha-(1,4)-, and alpha-(1,6)-glucosidic bonds. The present invention relates to mutant CC glucosylases (see AAB15178-B15184), which have increased CC thermostability, increased pH optimum and reduced isomaltose formation. The mutant proteins are useful for the selective production of glucose rather than alpha-1,6 linked disaccharide isomaltose

XX Sequence 616 AA;

Query Match 94.8%; Score 2628; DB 3; Length 616;
Best Local Similarity 99.6%; Pred. No. 6,5e-203;
Matches 504; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

QY 25 ATLDSWLSNEATVARTAILNNIGADGAWVGADSGIYVVASPTDNDPYFTWTRDSGLVLT 84
DB 1 ATLDSWLSNEATVARTAILNNIGADGAWVGADSGIYVVASPTDNDPYFTWTRDSGLVLT 60
QY 85 KTLVDLFRNGDTSILSTIENYISAQAIVQGISNPSGDLSSGAGLGEPEKFNVDERTAYTGSW 144
DB 61 KTLVDLFRNGDTSILSTIENYISAQAIVQGISNPSGDLSSGAGLGEPEKFNVDERTAYTGSW 120
QY 145 GRPQRDPALRATAMIGFGQMLDNGYTSTRTDVIWPLVNDLSYVAQYNNQGYDLMEE 204
DB 121 GRPQRDPALRATAMIGFGQMLDNGYTSTRTDVIWPLVNDLSYVAQYNNQGYDLMEE 180
QY 205 VNGSSFTTIAVOHRALVEGSAFATAVGSSGSCWCDSDAPEILCYQSFWTGSFILANFDS 264
DB 181 VNGSSFTTIAVOHRALVEGSAFATAVGSSGSCWCDSDAPEILCYQSFWTGSFILANFDS 240
QY 265 RSGKDANTLGSIHFPPEACDSTFQPCSPRALAHKKEVDSFRSITYLNDGLSSEA 324
DB 241 RSGKDANTLGSIHFPPEACDSTFQPCSPRALAHKKEVDSFRSITYLNDGLSSEA 300
QY 325 VAVGRYPEDTYNGNPMWFLCTLAABQLYDALYQMDKQGLSEVTVLDFPFKALYSDAAT 384
DB 301 VAVGRYPEDTYNGNPMWFLCTLAABQLYDALYQMDKQGLSEVTVLDFPFKALYSDAAT 360
QY 385 GTYSSSSSTYSIVDAVKTADGVSIVETHAASNGMSQYDKSDGEQLSARDLTWSYA 444
DB 361 GTYSSSSSTYSIVDAVKTADGVSIVETHAASNGMSQYDKSDGEQLSARDLTWSYA 420
QY 445 ALLTANNRRNSVVPASWGETSASSVPCTCATSAIGTYSSVTVTWSRISYATGTTTAT 504
DB 421 ALLTANNRRNSVVPASWGETSASSVPCTCATSAIGTYSSVTVTWSRISYATGTTTAT 480
QY 505 PTGSGSYTSTSKTTATASKTSTTSS 530
DB 481 PTGSGSYTSTSKTTATASKTSTTSS 506

```

RESULT 12

AAM55979 standard; protein; 616 AA.

AC AAM55979;

DT 27-JUL-1998 (first entry)

DE Aspergillus awamori glucosylase mutant S411A.

KM Aspergillus awamori; glucosylase; Aspergillus sp. mutant; fungal; food; fructose; corn; sweetener; 1,4-alpha-D-glucan glucosylase;

KW genetic engineering.
XX Synthetic.
OS Aspergillus awamori.
XX WO9803639-A1.
XX PN
XX PD 29-JAN-1998.
XX
XX 24-JUL-1997; 97WO-US012983.
XX
XX 24-JUL-1996; 96US-0022578P.
PR 02-AUG-1996; 96US-0023077P.
XX
XX (IOWA) UNIV IOWA STATE RES FOUND INC.
PA
XX Allen M, Fang T, Li Y, Liu H, Chen H, Coutinho P, Honzatko R,
PI Ford C;
XX WPI; 1998-120764/11.
XX
XX Genetically engineered fungal glucosylase - useful in, e.g. food
PT industry for production of high fructose corn sweeteners.
PS
PS Claim 10; Page: 97pp; English.
XX
XX The present sequence represents a specifically claimed mutant
CC glucosylase from Aspergillus awamori (1,4-alpha-D-glucan
CC glucosylase). The present invention describes fungal glucosylases
CC (FG) comprising: a mutation pair Asn20Cys coupled with Ala27Cys forming a
CC disulfide bond between the 2 members of the pair; and a 311-314loop or
CC Ser11Ala mutation. FG can be used in industry for the production of high
CC fructose corn sweeteners, while the glucose produced by glucosylase can
CC be crystallised or used in fermentation to produce organic products, e.g.
CC citric acid, ascorbic acid, lysine, glutamic acid or ethanol for
CC beverages and fuel. The mutations provide increased thermal stability,
CC reduced isomaltose formation and increased pH optimum. N.B. The present
CC sequence is not given in the specification but is derived from SEQ ID
CC NO.1 as stated in the claim
XX
SQ Sequence 616 AA;
Query Match 94.7%; Score 2625; DB 2; Length 616;
Best Local Similarity 99.4%; Pred. No. 1.1e-202;
Matches 503; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 25 ATLDSWLSNEATVARTAILNINIGADGAWVSGDSGIVVASTPSTNDPDTFTWTRDSGLVL 84
DB 1 ATLDSWLSNEATVARTAILNINIGADGAWVSGDSGIVVASTPSTNDPDTFTWTRDSGLVL 60
QY 85 KTLVULFRNGDTSLSTIENYISAQAIYVGISNPSGDLSSGGLGEPKFNDEMTYTGSM 144
DB 61 KTLVULFRNGDTSLSTIENYISAQAIYVGISNPSGDLSSGGLGEPKFNDEMTYTGSM 120
QY 145 GRPORDPALRTATAMIGFQOMLNDGYSTATDIWPLVRNLSVYAOVMNQTGDIWEE 204
DB 121 GRPORDPALRTATAMIGFQOMLNDGYSTATDIWPLVRNLSVYAOVMNQTGDIWEE 180
QY 205 VNGSSFFITAVOHRALVEGSAFAFAVGS8CWCDSQAPILLCYLOSFWTGSITLANPDS 264
DB 181 VNGSSFFITAVOHRALVEGSAFAFAVGS8CWCDSQAPILLCYLOSFWTGSITLANPDS 240
QY 265 RSGKANTLLGSIHFDPDEAACDDSTFQPCSPRALANHEKVVDSFRSITTLNDGSDSEA 324
DB 241 RSGKANTLLGSIHFDPDEAACDDSTFQPCSPRALANHEKVVDSFRSITTLNDGSDSEA 300
QY 325 VAVGVPEDTYNGPMFLCTLAABEOLYDALYOMDKOSLEVTQVSLDFPKALYSDAT 384
DB 301 VAVGVPEDTYNGPMFLCTLAABEOLYDALYOMDKOSLEVTQVSLDFPKALYSDAT 360
QY 385 GTYSSSSSTYSIVDAVKTFFADGFSIVETHAASNGSMEQYKSDGEOLASRDLTWSYA 444
DB 361 GTYSSSSSTYSIVDAVKTFFADGFSIVETHAASNGSMEQYKSDGEOLASRDLTWSYA 420

QY 445 ALLTANNRNSVVPASWGETSASSVPGCAATSAIGTSSVTVTSPSTIVATGGTTTAT 504
DB 421 ALLTANNRNSVVPASWGETSASSVPGCAATSAIGTSSVTVTSPSTIVATGGTTTAT 480
QY 505 PTGSGSVTSTSKTATATASKTSTTSS 530
DB 481 PTGSGSVTSTSKTATATASKTSTTSS 506

RESULT 13
AAB15180
ID AAB15180 standard; protein; 616 AA.
XX
XX AAB15180;
XX
XX 11-DEC-2000 (first entry)
XX
XX Aspergillus awamori mutant glucosylase S411A substitution.
DE
XX Glucosylase; enzyme; carbohydrase; glucose;
KW 1,4-alpha-D-glucan glucosylase; muten; mutation.
XX
XX Aspergillus awamori.
OS Synthetic.
OS
FH Key Location/Qualifiers
FT Misc-difference 411
FT W0200043504-A1.
XX
XX 27-JUL-2000.
XX
XX 10-JAN-2000; 2000WO-US000532.
XX
XX 22-JAN-1999; 99US-00236063.
PR
XX (IOWA) UNIV IOWA STATE RES FOUND INC.
PA
XX Allen MJ, Fang T, Li Y, Liu H, Chen H, Coutinho P, Honzatko R,
PI Ford C;
XX WPI; 2000-514725/46.
XX
XX Fungal glucosylase for selective production of glucose rather than alpha
PT -1,6 linked disaccharide isomaltose, has mutation pair Asn20Cys coupled
PT with Ala27Cys forming disulfide bond between the two stabilizing members.
XX
PS Claim 10; Page: 160pp; English.
XX
XX Glucosylase (1,4-alpha-D-glucan glucosylase; E.C. 3.2.1.3) is a
CC carbohydrase. This enzyme cleaves D-glucose from the nonreducing ends of
CC maltooligosaccharides, attacking alpha-(1,4)-, and alpha-(1,6)-glucosidic
CC bonds. The present invention relates to mutant glucosylases, which have
CC increased thermostability, increased pH optimum and reduced isomaltose
CC formation. The mutant proteins are useful for the selective production of
CC glucose rather than alpha-1,6 linked disaccharide isomaltose. The present
CC sequence is one such mutant enzyme. Note: The present sequence is not
CC shown in the specification but is derived from the Aspergillus awamori
CC wild-type glucosylase sequence given in pages 152-153 of the sequence
CC listing (SEQ ID 1)
XX
SQ Sequence 616 AA;
Query Match 94.7%; Score 2625; DB 3; Length 616;
Best Local Similarity 99.4%; Pred. No. 1.1e-202;
Matches 503; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 25 ATLDSWLSNEATVARTAILNINIGADGAWVSGDSGIVVASTPSTNDPDTFTWTRDSGLVL 84
DB 1 ATLDSWLSNEATVARTAILNINIGADGAWVSGDSGIVVASTPSTNDPDTFTWTRDSGLVL 60

435
Not claimed

```

QY 85 KTLVDLFRNGDTSLSLTIENTYISAQAIYVQGISNPSGDLSSGAGLGEPEKFNVDETAYTGSW 144
DB 61 KTLVDLFRNGDTSLSLTIENTYISAQAIYVQGISNPSGDLSSGAGLGEPEKFNVDETAYTGSW 120
QY 145 GRPORDGPALRATAMIGFGQWLDNGYTSATDIWPLVRNDLSYVAQYNNQGYDLMEE 204
DB 121 GRPORDGPALRATAMIGFGQWLDNGYTSATDIWPLVRNDLSYVAQYNNQGYDLMEE 180
QY 205 VNGSSPFTIAVOHRAIVEGSAFATAVAGSSCSWCDSQAPELICYSQSFMTGSFILLANPDSS 264
DB 181 VNGSSPFTIAVOHRAIVEGSAFATAVAGSSCSWCDSQAPELICYSQSFMTGSFILLANPDSS 240
QY 265 RSGGDANTLLGSIHTFPDEAACDDSTFQPCSPRALANHKEVVDSPRSIYTLNDGLSDSEA 324
DB 241 RSGGDANTLLGSIHTFPDEAACDDSTFQPCSPRALANHKEVVDSPRSIYTLNDGLSDSEA 300
QY 325 VAVGRYEDDTYYNNPWFCTLLAAAEOLYDALYQMDKQGLEVTVDVSLDFPKALYSDAAT 384
DB 301 VAVGRYEDDTYYNNPWFCTLLAAAEOLYDALYQMDKQGLEVTVDVSLDFPKALYSDAAT 360
QY 385 GTYSSSSSTYSIYDAVKTFPADGFVSIIVETHAANGSMSEQYDKSDGEQLSARDLTWSYA 444
DB 361 GTYSSSSSTYSIYDAVKTFPADGFVSIIVETHAANGSMSEQYDKSDGEQLSARDLTWSYA 420
QY 445 ALLTANNRRNSVVPASWGETSASSVPGTCAATSAIGTYSVTVTWSPIVATGGTITTTAT 504
DB 421 ALLTANNRRNSVVPASWGETSASSVPGTCAATSAIGTYSVTVTWSPIVATGGTITTTAT 480
QY 505 PTSGGSVTSKTTATASKTSTTRS 530
DB 481 PTSGGSVTSKTTATASKTSTTRS 506

```

RESULT 14

AAM55976
ID AAM55976 standard; protein: 616 AA.

AC AAM55976;

DT 27-JUL-1998 (first entry)

DE Aspergillus awamori glucoamylase mutant N20C, A27C.

KM Aspergillus awamori; glucoamylase; Aspergillus sp; mutant; fungal; food;
glucose; corn; sweetener; 1,4-alpha-D-glucan glucohydrolase;

KM genetic engineering.

OS Synthetic.

OS Aspergillus awamori.

FT Key Location/Qualifiers

FT Disulfide-bond 20..27

PN WO9803639-A1.

PD 29-JAN-1998.

PF 24-JUL-1997; 97WO-US012983.

PR 24-JUL-1996; 96US-0022578P.

PR 02-AUG-1996; 96US-0023077P.

PA (IOWA) UNIV IOWA STATE RES FOUND INC.

PI Allen M, Fang T, Li Y, Liu H, Chen H, Coutinho P, Honzatho R,
Ford C;

DR WPI; 1998-120764/11.

XX Genetically engineered fungal glucoamylase - useful in, e.g. food

PT industry for production of high fructose corn sweeteners.

PS Claim 1; Page: 97pp; English.

XX The present sequence represents a specifically claimed mutant
CC glucoamylase from Aspergillus awamori (1,4-alpha-D-glucan
CC glucohydrolase). The present invention describes fungal glucoamylases
CC (PG) comprising: a mutation pair Asn20Cys coupled with Ala27Cys forming a
CC disulphide bond between the 2 members of the pair; and a 311-314loop or
CC Ser411Ala mutation. PG can be used in industry for the production of high
CC fructose corn sweeteners, while the glucose produced by glucoamylase can
CC be crystallised or used in fermentation to produce organic products, e.g.
CC citric acid, ascorbic acid, lysine, glutamic acid or ethanol for
CC beverages and fuel. The mutations provide increased thermal stability.
CC reduced isomaltose formation and increased pH optimum. N.B. The present
CC sequence is not given in the specification but is derived from SEQ ID
CC No:1 as stated in the claim

SO Sequence 616 AA;

Query Match 94.4%; Score 2615; DB 2; Length 616;
Best Local Similarity 99.2%; Pred. No. 7.3e-202;
Matches 502; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 25 ATLDMSLNEATYARATILNNGADGAWGADSGIYVASTDNPYFTYMTDRDGLVL 84
DB 1 ATLDMSLNEATYARATILNNGADGAWGADSGIYVASTDNPYFTYMTDRDGLVL 60

QY 85 KTLVDLFRNGDTSLSLTIENTYISAQAIYVQGISNPSGDLSSGAGLGEPEKFNVDETAYTGSW 144
DB 61 KTLVDLFRNGDTSLSLTIENTYISAQAIYVQGISNPSGDLSSGAGLGEPEKFNVDETAYTGSW 120

QY 145 GRPORDGPALRATAMIGFGQWLDNGYTSATDIWPLVRNDLSYVAQYNNQGYDLMEE 204
DB 121 GRPORDGPALRATAMIGFGQWLDNGYTSATDIWPLVRNDLSYVAQYNNQGYDLMEE 180

QY 205 VNGSSPFTIAVOHRAIVEGSAFATAVAGSSCSWCDSQAPELICYSQSFMTGSFILLANPDSS 264
DB 181 VNGSSPFTIAVOHRAIVEGSAFATAVAGSSCSWCDSQAPELICYSQSFMTGSFILLANPDSS 240

QY 265 RSGGDANTLLGSIHTFPDEAACDDSTFQPCSPRALANHKEVVDSPRSIYTLNDGLSDSEA 324
DB 241 RSGGDANTLLGSIHTFPDEAACDDSTFQPCSPRALANHKEVVDSPRSIYTLNDGLSDSEA 300

QY 325 VAVGRYEDDTYYNNPWFCTLLAAAEOLYDALYQMDKQGLEVTVDVSLDFPKALYSDAAT 384
DB 301 VAVGRYEDDTYYNNPWFCTLLAAAEOLYDALYQMDKQGLEVTVDVSLDFPKALYSDAAT 360

QY 385 GTYSSSSSTYSIYDAVKTFPADGFVSIIVETHAANGSMSEQYDKSDGEQLSARDLTWSYA 444
DB 361 GTYSSSSSTYSIYDAVKTFPADGFVSIIVETHAANGSMSEQYDKSDGEQLSARDLTWSYA 420

QY 445 ALLTANNRRNSVVPASWGETSASSVPGTCAATSAIGTYSVTVTWSPIVATGGTITTTAT 504
DB 421 ALLTANNRRNSVVPASWGETSASSVPGTCAATSAIGTYSVTVTWSPIVATGGTITTTAT 480

QY 505 PTSGGSVTSKTTATASKTSTTRS 530
DB 481 PTSGGSVTSKTTATASKTSTTRS 506

RESULT 15

AAB15178
ID AAB15178 standard; protein: 616 AA.

AC AAB15178;

DT 11-DEC-2000 (first entry)

DE Aspergillus awamori mutant glucoamylase N20C/A27C substitution.

XX Glucoamylase; enzyme; carboxydase; glucose;

KM 1,4-alpha-D-glucan glucohydrolase; mutcin; mutation.

OS Aspergillus awamori.

OS Synthetic.

XX Key Location/Qualifiers
 FH Disulfide-bond 20..27
 FT Misc-difference 20
 FT Misc-difference 27 /note= "wild-type Asn substituted by Cys"
 FT Misc-difference 27 /note= "wild-type Ala substituted by Cys"
 PN MO200043504-A1.
 PD 27-JUL-2000.
 XX 10-JAN-2000; 2000WO-US000532.
 XX 22-JAN-1999; 99US-00236063.
 XX (IOWA) UNIV IOWA STATE RES FOUND INC.
 XX Allen MJ, Fang T, Li Y, Liu H, Chen H, Coutinho P, Honzatko R;
 PI Ford C;
 XX WPI; 2000-514725/46.
 XX
 PT Fungal glucosylase for selective production of glucose rather than alpha
 PT -1,6 linked disaccharide isomaltose, has mutation pair Asn20Cys coupled
 PT with Ala27Cys forming disulfide bond between the two stabilizing members.
 XX
 PS Claim 1; Page; 160pp; English.
 XX
 CC Glucosylase (1,4-alpha-D-glucan glucosylase; E.C. 3.2.1.3) is a
 CC carbohydriase. This enzyme cleaves D-glucose from the nonreducing ends of
 CC maltooligosaccharides, attacking alpha-(1,4)-, and alpha-(1,6)-glucosidic
 CC bonds. The present invention relates to mutant glucosylases, which have
 CC increased thermostability, increased pH optimum and reduced isomaltose
 CC formation. The mutant proteins are useful for the selective production of
 CC glucose rather than alpha-1,6 linked disaccharide isomaltose. The present
 CC sequence is one such mutant enzyme. Note: The present sequence is not
 CC shown in the specification but is derived from the Aspergillus awamori
 CC wild-type glucosylase sequence given in pages 152-153 of the sequence
 CC listing (Seq ID 1)
 CC
 SQ Sequence 616 AA;
 Query Match 94.4%; Score 2615; DB 3; Length 616;
 Best Local Similarity 99.2%; Pred. No. 7.3e-202;
 Matches 502; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 25 ATLDSMLNSNEATVARTAILNINIGADGAWSGADSGIVVSPSTNDPDIYFTWTRDSGLVL 84
 DB 1 ATLDSMLNSNEATVARTAILNINIGADGAWSGADSGIVVSPSTNDPDIYFTWTRDSGLVL 60
 QY 85 KTLVDLFRNGDSTSLSTIBENYISAQAIIVGIGISNPSGDLSSGAGLGEPEKFNDETAVTGSM 144
 DB 61 KTLVDLFRNGDSTSLSTIBENYISAQAIIVGIGISNPSGDLSSGAGLGEPEKFNDETAVTGSM 120
 QY 145 GRPORDGPRALRTAMIGFGOMLLDNGYSTATDIWPIVRNDLSYVAQVWNOTGYDLWE 204
 DB 121 GRPORDGPRALRTAMIGFGOMLLDNGYSTATDIWPIVRNDLSYVAQVWNOTGYDLWE 180
 QY 205 VNGSSFFITIAVQHRALVEGSAFATVAGSSCSWCDSDQAPILCYLOSFWTGSFILANFDS 264
 DB 181 VNGSSFFITIAVQHRALVEGSAFATVAGSSCSWCDSDQAPILCYLOSFWTGSFILANFDS 240
 QY 265 RSGKDANTILGSIHTFPDEAACDSTFQCSPRALANHKEVVDSPRSIYTLNDGLSDSEA 324
 DB 241 RSGKDANTILGSIHTFPDEAACDSTFQCSPRALANHKEVVDSPRSIYTLNDGLSDSEA 300
 QY 325 VAVGRYPEDTYNGNPMFLCTLAABQLYDALYQWDKQSLFVTVSLDFEALYSDAAT 384
 DB 301 VAVGRYPEDTYNGNPMFLCTLAABQLYDALYQWDKQSLFVTVSLDFEALYSDAAT 360
 QY 385 GTYSSSSSTYSIVAVKTFADGPFVIVETHAASNGSMSEQYDKSDGEQLSARDLTWSYA 444

DB 361 GTYSSSSSTYSIVAVKTFADGPFVIVETHAASNGSMSEQYDKSDGEQLSARDLTWSYA 420
 QY 445 ALLTANNRRNSVPAKMGETGASSVPGCATSAIGTYSVTYTSWPSIVATGCTTTAT 504
 DB 421 ALLTANNRRNSVPAKMGETGASSVPGCATSAIGTYSVTYTSWPSIVATGCTTTAT 480
 QY 505 PTGSGSVTSTSKTATATASKTSTTTS 530
 DB 481 PTGSGSVTSTSKTATATASKTSTTTS 506

Search completed: June 17, 2004, 17:45:59
 Job time : 62 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2004, 17:47:36 ; Search time 50 Seconds
(without alignments)
3015.110 Million cell updates/sec

Title: US-10-038-723-2
Perfect score: 2771
Sequence: 1 MSFRLSLSGVCTGLANV.....SKTTATSKTSTTTSGMSL 534

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1163542 seqs, 282313646 residues
Total number of hits satisfying chosen parameters: 1163542

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2771	100.0	534	9	US-09-908-395-2
2	2771	100.0	534	10	US-09-821-616-9
3	2771	100.0	534	14	US-10-038-723-2
4	2771	100.0	534	15	US-10-421-586-2
5	2771	100.0	534	15	US-10-421-586-3
6	2742	99.0	640	14	US-10-038-723-13
7	2742	99.0	640	15	US-10-421-586-14
8	2735	99.7	743	15	US-10-418-836-9
9	2735	99.7	763	15	US-10-418-836-13
10	2735	99.7	979	15	US-10-418-836-10
11	2735	99.7	979	15	US-10-418-836-16
12	2720.5	98.2	738	15	US-10-418-836-19
13	2716	98.0	741	15	US-10-418-836-30
14	2715.5	98.0	972	15	US-10-418-836-38
15	2715	98.0	739	15	US-10-418-836-26

16	2715	98.0	739	15	US-10-418-836-27	Sequence 27, Appl
17	2714.5	98.0	740	15	US-10-418-836-28	Sequence 28, Appl
18	2714.5	98.0	740	15	US-10-418-836-34	Sequence 34, Appl
19	2714.5	98.0	742	15	US-10-418-836-29	Sequence 29, Appl
20	2714	97.9	743	15	US-10-418-836-31	Sequence 31, Appl
21	2714	97.9	743	15	US-10-418-836-35	Sequence 35, Appl
22	2714	97.9	743	15	US-10-418-836-39	Sequence 39, Appl
23	2523	91.1	1095	14	US-10-228-063-45	Sequence 45, Appl
24	1850.5	66.8	631	14	US-10-213-990-54	Sequence 54, Appl
25	1689	61.0	618	10	US-09-821-616-74	Sequence 34, Appl
26	1674.5	60.4	591	10	US-09-821-616-7	Sequence 7, Appl
27	1548.5	55.9	647	15	US-10-369-493-1109	Sequence 3109, Appl
28	1425.5	51.4	581	14	US-10-281-673-3	Sequence 3, Appl
29	1408.5	50.8	704	14	US-10-213-990-21	Sequence 21, Appl
30	1380.5	49.8	620	14	US-10-213-990-33	Sequence 33, Appl
31	1250	45.1	441	15	US-10-369-493-12402	Sequence 12402, A
32	1132	40.9	432	15	US-10-369-493-12533	Sequence 12533, A
33	781.5	28.2	450	15	US-10-369-493-2490	Sequence 2490, Ap
34	725	26.2	579	14	US-10-228-063-49	Sequence 49, Appl
35	597	21.5	549	15	US-10-369-493-22096	Sequence 22096, A
36	164.5	5.9	2344	9	US-09-815-242-12713	Sequence 12713, A
37	160	5.8	35	10	US-09-821-616-4	Sequence 4, Appl
38	159.5	5.8	2283	14	US-10-172-502-4	Sequence 4, Appl
39	159	5.7	2271	12	US-10-282-122A-43924	Sequence 43924, A
40	158	5.7	1283	15	US-10-369-493-22616	Sequence 22616, A
41	155.5	5.6	1621	14	US-10-185-990-10	Sequence 10, Appl
42	155.5	5.6	1626	14	US-10-185-990-11	Sequence 11, Appl
43	154.5	5.6	1031	9	US-09-815-242-10932	Sequence 10932, A
44	152.5	5.5	2117	15	US-10-120-801-63	Sequence 63, Appl
45	152	5.5	45	14	US-10-228-063-53	Sequence 53, Appl

ALIGNMENTS

RESULT 1
US-09-908-395-2
Sequence 2, Appli
Patent No. US20020164723A1
GENERAL INFORMATION:
APPLICANT: liaw, Gai
APPLICANT: Pedersen, Sven
TITLE OF INVENTION: A Method of Producing Saccharide
FILE REFERENCE: 5318.200-US
CURRENT APPLICATION NUMBER: US/09/908.395
CURRENT FILING DATE: 2001-07-18
PRIOR FILING DATE: PRIOR FILING DATE: 1998-11-23
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 2
LENGTH: 534
TYPE: PRT
ORGANISM: Aspergillus Niger
FEATURE:
NAME/KEY: SIGNAL
LOCATION: (1)...(24)
US-09-908-395-2

Query Match 100.0% Score 2771; DB 9; Length 534;
Best Local Similarity 100.0% Pred. No. 2.7e-237;
Matches 534; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSFRLSLSGVCTGLANVSKATLDSMTSNEATVARTAILNNIGAGAVSGADSGI 60
DB 1 MSFRLSLSGVCTGLANVSKATLDSMTSNEATVARTAILNNIGAGAVSGADSGI 60
QY 61 VVASPSTNDPYPYTWTRDSGLVKTIVDLFRNGDTSLSTIENISAQAIVOGISNPSG 120
DB 61 VVASPSTNDPYPYTWTRDSGLVKTIVDLFRNGDTSLSTIENISAQAIVOGISNPSG 120

Post-processed
-a-prv-cdm

```
QY 121 DSSGAGLGPKNVDETAATYGSWGRPQDGPALRATAMIGFGQWLLDNGYTSTATDIW 180
| | | | |
DB 121 DSSGAGLGPKNVDETAATYGSWGRPQDGPALRATAMIGFGQWLLDNGYTSTATDIW 180
QY 181 PLVNRNDLSVYAAQYWNQGYDLMEEVNGSSPFTTAVQHRALVEGSAFATAVGSSCSWCDSQ 240
| | | | |
DB 181 PLVNRNDLSVYAAQYWNQGYDLMEEVNGSSPFTTAVQHRALVEGSAFATAVGSSCSWCDSQ 240
QY 241 APEILCYLQSFMTGSPFLIANFDSRSRSGKANTLLGSIHTPDEAACDSTFQPCSPRALA 300
| | | | |
DB 241 APEILCYLQSFMTGSPFLIANFDSRSRSGKANTLLGSIHTPDEAACDSTFQPCSPRALA 300
QY 301 NHKEVDSFRSIIYTLNDGLSDSEAVAGRYPEDTYYNGNPFLLCTTAAEQLYDALYQMD 360
| | | | |
DB 301 NHKEVDSFRSIIYTLNDGLSDSEAVAGRYPEDTYYNGNPFLLCTTAAEQLYDALYQMD 360
QY 361 KQGSLEVTDVSLDPEFKALYSDAATGTYSSSSSTYSIVDAVKTPADGFSIVETHAASNG 420
| | | | |
DB 361 KQGSLEVTDVSLDPEFKALYSDAATGTYSSSSSTYSIVDAVKTPADGFSIVETHAASNG 420
QY 421 SMSEQYDKSDGEQLSARDLTWSYAALLTANNRRNSVVPASWGETSASVPOTCAATSAIG 480
| | | | |
DB 421 SMSEQYDKSDGEQLSARDLTWSYAALLTANNRRNSVVPASWGETSASVPOTCAATSAIG 480
QY 481 TYSSVTVTSWPSIYATGTTTATPTGSGSVTSTSKTTATASKTSTTRSGMSL 534
| | | | |
DB 481 TYSSVTVTSWPSIYATGTTTATPTGSGSVTSTSKTTATASKTSTTRSGMSL 534
```

```
RESULT 2
US-09-821-616-9
; Sequence 9, Application US/09821616
; Publication No. US20030027290A1
; GENERAL INFORMATION:
; APPLICANT: Nielsen, Bjarne R.
; APPLICANT: Lembeck, Jan
; TITLE OF INVENTION: Thermostable Glucoamylase
; FILE REFERENCE: 5279, 200-US
; CURRENT APPLICATION NUMBER: US/09/821, 616
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/199, 290
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-11-24
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 1557/97
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-12-30
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 0925/98
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-07-10
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/070, 746
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-01-08
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/094, 344
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-07-28
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/979, 673
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/107, 657
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-30
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 9
; LENGTH: 534
; TYPE: PRT
; ORGANISM: Aspergillus niger
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)...(24)
US-09-821-616-9
```

```
Query Match 100.0%; Score 2771; DB 10; Length 534;
Best Local Similarity 100.0%; Pred. No. 2,7e-237;
Matches 534; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSFSLALSLGLVCTGLANVISKRAATLDSWLSNEATYAKTALINNIGADGAWVSGANSGI 60
| | | | |
DB 1 MSFSLALSLGLVCTGLANVISKRAATLDSWLSNEATYAKTALINNIGADGAWVSGANSGI 60
```

```
QY 61 VVASPSTDNPDYFYTWTRDSGLVLTLYDLFRNGDTSLSTIENYISAQAIYQGISNPSG 120
| | | | |
DB 61 VVASPSTDNPDYFYTWTRDSGLVLTLYDLFRNGDTSLSTIENYISAQAIYQGISNPSG 120
QY 121 DSSGAGLGPKNVDETAATYGSWGRPQDGPALRATAMIGFGQWLLDNGYTSTATDIW 180
| | | | |
DB 121 DSSGAGLGPKNVDETAATYGSWGRPQDGPALRATAMIGFGQWLLDNGYTSTATDIW 180
QY 181 PLVNRNDLSVYAAQYWNQGYDLMEEVNGSSPFTTAVQHRALVEGSAFATAVGSSCSWCDSQ 240
| | | | |
DB 181 PLVNRNDLSVYAAQYWNQGYDLMEEVNGSSPFTTAVQHRALVEGSAFATAVGSSCSWCDSQ 240
QY 241 APEILCYLQSFMTGSPFLIANFDSRSRSGKANTLLGSIHTPDEAACDSTFQPCSPRALA 300
| | | | |
DB 241 APEILCYLQSFMTGSPFLIANFDSRSRSGKANTLLGSIHTPDEAACDSTFQPCSPRALA 300
QY 301 NHKEVDSFRSIIYTLNDGLSDSEAVAGRYPEDTYYNGNPFLLCTTAAEQLYDALYQMD 360
| | | | |
DB 301 NHKEVDSFRSIIYTLNDGLSDSEAVAGRYPEDTYYNGNPFLLCTTAAEQLYDALYQMD 360
QY 361 KQGSLEVTDVSLDPEFKALYSDAATGTYSSSSSTYSIVDAVKTPADGFSIVETHAASNG 420
| | | | |
DB 361 KQGSLEVTDVSLDPEFKALYSDAATGTYSSSSSTYSIVDAVKTPADGFSIVETHAASNG 420
QY 421 SMSEQYDKSDGEQLSARDLTWSYAALLTANNRRNSVVPASWGETSASVPOTCAATSAIG 480
| | | | |
DB 421 SMSEQYDKSDGEQLSARDLTWSYAALLTANNRRNSVVPASWGETSASVPOTCAATSAIG 480
QY 481 TYSSVTVTSWPSIYATGTTTATPTGSGSVTSTSKTTATASKTSTTRSGMSL 534
| | | | |
DB 481 TYSSVTVTSWPSIYATGTTTATPTGSGSVTSTSKTTATASKTSTTRSGMSL 534
```

```
RESULT 3
US-10-038-723-2
; Sequence 2, Application US/10038723
; Publication No. US20030032163A1
; GENERAL INFORMATION:
; APPLICANT: Nielsen, Bjarne Roenfeldt
; APPLICANT: Pedersen, Henrik
; APPLICANT: Vind, Jesper
; APPLICANT: Hendiksen, Hanne Vang
; APPLICANT: Frandsen, Torben Peter
; TITLE OF INVENTION: Glucoamylase Variants
; FILE REFERENCE: 5636, 200-US
; CURRENT APPLICATION NUMBER: US/10/038, 723
; PRIOR FILING DATE: 2002-01-02
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/351, 814
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-07-12
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: PA 1998 00937
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: PA 1998 01667
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-12-17
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/093, 528
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-07-21
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/115, 545
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-01-12
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 534
; TYPE: PRT
; ORGANISM: Aspergillus niger
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)...(24)
US-10-038-723-2
```

```
Query Match 100.0%; Score 2771; DB 14; Length 534;
Best Local Similarity 100.0%; Pred. No. 2,7e-237;
Matches 534; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy	1	MSFRSLIALSGVLCGLANVISRATLDSLSNEATVAPRAIINNNGADAMVSGADSGI	60
Dd	1	MSFRSLIALSGVLCGLANVISRATLDSLSNEATVAPRAIINNNGADAMVSGASGI	60
Qy	61	VVASPTSDPDFFYTTWRDSGLVTKTLVDLFRNGDPSLSTIENTYISAQAIVOGISNPSG	120
Dd	61	VVASPTSDPDFFYTTWRDSGLVTKTLVDLFRNGDPSLSTIENTYISAQAIVOGISNPSG	120
Qy	121	DLSSGAGLGEPRKNVDETAVTGSGMGRPORDGAPRALATAMIGFQOMLNDNGYSTDADIW	180
Dd	121	DLSSGAGLGEPRKNVDETAVTGSGMGRPORDGAPRALATAMIGFQOMLNDNGYSTDADIW	180
Qy	181	PLVRNDLSVVAQYMNQGYDLMEENVGSSFFTIYVQHRALVEGSAFATAVGSSCSWCDSQ	240
Dd	181	PLVRNDLSVVAQYMNQGYDLMEENVGSSFFTIYVQHRALVEGSAFATAVGSSCSWCDSQ	240
Qy	241	APELLICYLOSFWTGSFIILANFDSRSGKXANTLLGSIHTFDPEAACDSTFFQCSPPALA	300
Dd	241	APELLICYLOSFWTGSFIILANFDSRSGKXANTLLGSIHTFDPEAACDSTFFQCSPPALA	300
Qy	301	NHKEVVDSPRSYITLNDGSDSEBAVAVGRPEPTYNNGNMFECTJLAAAEQLDALYQWD	360
Dd	301	NHKEVVDSPRSYITLNDGSDSEBAVAVGRPEPTYNNGNMFECTJLAAAEQLDALYQWD	360
Qy	361	KQGSLEVTDVSLDFFKALYSDAATGYSSSSSSTYSSIIVDAVKTFADGFVSIIVETHAASNG	420
Dd	361	KQGSLEVTDVSLDFFKALYSDAATGYSSSSSSTYSSIIVDAVKTFADGFVSIIVETHAASNG	420
Qy	421	SMSBOYDKSDGEQLSARDLTMSYAALLTANNRRNSVVPASWGETSASSVPGTCAATSAIG	480
Dd	421	SMSBOYDKSDGEQLSARDLTMSYAALLTANNRRNSVVPASWGETSASSVPGTCAATSAIG	480
Qy	481	TYSSVTLVSWBEIVATGCTTTATPFGGSSVYSTKTTATASTKSTTTTRGSMGL	534
Dd	481	TYSSVTLVSWBEIVATGCTTTATPFGGSSVYSTKTTATASTKSTTTTRGSMGL	534

Dehpari
28

```

US-10-421-586-2
RESULT 4
US-10-421-586-2
; Sequence 2, Application US/10421586
; Publication No. US20040002142A1
; GENERAL INFORMATION:
; APPLICANT: Melsen, Bjarne Ronfeldt
; APPLICANT: Svendsen, Allan
; APPLICANT: Pedersen, Henrik
; APPLICANT: Vind, Jesper
; APPLICANT: Hendriksen, Hanne Vang
; APPLICANT: Frandsen, Torben Peter
; TITLE OF INVENTION: Glucosylase Variants
; FILE REFERENCE: 5967, 210-US
; CURRENT APPLICATION NUMBER: US/10/421, 586
; CURRENT FILING DATE: 2003-04-23
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 534
; TYPE: prt
; ORGANISM: Aspergillus niger
US-10-421-586-2

```

Query Match	Similarity	Score	DB	Length
Best Local	100.0%	2771	15	534
Match	534	Conservative	0	Mismatch
			0	Indels
			0	Gaps
			0	

Qy	121	LLSSAGJGEPEFNDETAAYGSGWRPORDDGALATAMIGGOMLLDNGYSTADIYW	180
Db	121	DLSSAGJGEPEFNDETAAYGSGWRPORDDGALATAMIGGOMLLDNGYSTADIYW	180
Qy	181	PLVRNDLSYVAQYWNQYGTJDLMEEWNGSSFFTIAVQHRALVEGSAFATAVGSSCWCDSQ	240
Db	181	PLVRNDLSYVAQYWNQYGTJDLMEEWNGSSFFTIAVQHRALVEGSAFATAVGSSCWCDSQ	240
Qy	241	APELLICYIQSWTGSFIILANFDSSNSGKDANTLLSIRHFDPEAACDSDTFQPCSPRALA	300
Db	241	APELLICYIQSWTGSFIILANFDSSNSGKDANTLLSIRHFDPEAACDSDTFQPCSPRALA	300
Qy	301	NHKEVVDSEFRSIYTLNDGLDSEBEAVALGRYPEDTYNNGNPNFLCTAAAEOLYDALYQMD	360
Db	301	NHKEVVDSEFRSIYTLNDGLDSEBEAVALGRYPEDTYNNGNPNFLCTAAAEOLYDALYQMD	360
Qy	361	KQGSLEVDVSLDPEFKALYSDPATGYSSSSSTYGSYDAVATFPADGFVSIYETHAASNG	420
Db	361	KQGSLEVDVSLDPEFKALYSDPATGYSSSSSTYGSYDAVATFPADGFVSIYETHAASNG	420
Qy	421	SMSBEYDKSDGEOULARDLTWSYALLPLANNRNRNSVVPASWETSASSVPGTCAATSALG	480
Db	421	SMSBEYDKSDGEOULARDLTWSYALLPLANNRNRNSVVPASWETSASSVPGTCAATSALG	480
Qy	481	TYSSVTWVSWBSIYATGGTTTATPAGTGSVSTSKTATPASTKSTTTTSSGMSL	534
Db	481	TYSSVTWVSWBSIYATGGTTTATPAGTGSVSTSKTATPASTKSTTTTSSGMSL	534

```

RESULT 5
US-10-421-586-3
: Sequence 3, Application US/10421586
: Publication No. US20040002142A1
: GENERAL INFORMATION:
: APPLICANT: Nielsen, Bjarne Ronfeldt
: APPLICANT: Svendsen, Allan
: APPLICANT: Pedersen, Henrik
: APPLICANT: Vind, Jesper
: APPLICANT: Hendriksen, Haane Vang
: APPLICANT: Frandsen, Torben Peter
: TITLE OF INVENTION: Glucosylase Variants
: FILE REFERENCE: 5667,210-US
: CURRENT APPLICATION NUMBER: US/10/421,586
: NUMBER OF SEQ ID NOS: 18
: SOFTWARE: PatentIn version 3.2
: SEQ ID NO 3
: LENGTH: 534
: TYPE: PRT
: ORGANISM: Aspergillus niger
: FEATURE:
: NAME/KEY: Signal
: LOCATION: (1)..(24)
: US-10-421-586-3

```

Query Match	100.0%;	Score 2771;	DB 15;	Length 534;
Best Local Similarity	100.0%;	Pred. No. 2.7e-237;		
Matches 534;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0

Qy	1	MSRSLALISGLVCTGLANVISKRATLDSMWLNKNTVARTILNNIGDAGMVGASDGI	60
Db	1	MSRSLALISGLVCTGLANVISKRATLDSMWLNKNTVARTILNNIGDAGMVGASDGI	60
Qy	61	VVASPSTNDPDYFYTWTRDSGLVLTVLDFRNGDTSLLSTIENYISAQAIVOGISNPSG	120
Db	61	VVASPSTNDPDYFYTWTRDSGLVLTVLDFRNGDTSLLSTIENYISAQAIVOGISNPSG	120
Qy	121	DLSSAGAGLEPPFNVDERTAYTSGWGRPORQDGPALRATAMIGFGOMLLDNGTSTATIIYW	180
Db	121	DLSSAGAGLEPPFNVDERTAYTSGWGRPORQDGPALRATAMIGFGOMLLDNGTSTATIIYW	180
Qy	181	PLVRNDLSVYAQYNQGTGDLMEEVNGSSPFTIAYQHRALVEGSAFATANGSSCSKWDQ	240
Db	181	PLVRNDLSVYAQYNQGTGDLMEEVNGSSPFTIAYQHRALVEGSAFATANGSSCSKWDQ	240

```
Db      181 PLVNDLSYYAQAQWNOGTGYDLMEBVNGSSPFTTAVQHRALVEGSAFATAVGSSCSCWCDQ 240
Qy      241 APEILCYLGQFWTGSFTLIANFDSRSRSGKDNANTLLGSIHTFDPEAACDSTFQPCSPRALA 300
Db      241 APEILCYLGQFWTGSFTLIANFDSRSRSGKDNANTLLGSIHTFDPEAACDSTFQPCSPRALA 300
Qy      301 NHKEVVDSPRSIYTLNDGLSDSEAVAVGRYPEDTYNGNPMWFLCTTAAAEQLYDALYQMD 360
Db      301 NHKEVVDSPRSIYTLNDGLSDSEAVAVGRYPEDTYNGNPMWFLCTTAAAEQLYDALYQMD 360
Qy      361 KQSLFETDVSLDPEFKALYSDAATGCTYSSSSSTYSIYDAVKTPADGFVSIIVETHAASNG 420
Db      361 KQSLFETDVSLDPEFKALYSDAATGCTYSSSSSTYSIYDAVKTPADGFVSIIVETHAASNG 420
Qy      421 SMSEQYKSDGEOLASDILTWSYAALLTANNRRNSVVPASWGETSASVPCTCAATSAIG 480
Db      421 SMSEQYKSDGEOLASDILTWSYAALLTANNRRNSVVPASWGETSASVPCTCAATSAIG 480
Qy      481 TYSSTVTSWPSIYATGCTTTTATPTGSGSVTSTSKTTATATASKTSTTTRSGMSL 534
Db      481 TYSSTVTSWPSIYATGCTTTTATPTGSGSVTSTSKTTATATASKTSTTTRSGMSL 534
```

RESULT 6
US-10-038-723-13

Sequence 13, Application US/10038723
Publication No. US20030032163A1

GENERAL INFORMATION:

APPLICANT: Nielsen, Bjarne Roenfeldt

APPLICANT: Svendsen, Allan

APPLICANT: Pedersen, Henrik

APPLICANT: Vind, Jesper

APPLICANT: Hendriksen, Hanne Vang

APPLICANT: Frandsen, Torben Peter

TITLE OR INVENTION: Glucoamylase Variants

FILE REFERENCE: 5636.200-US

CURRENT APPLICATION NUMBER: US/10/038, 723

CURRENT FILING DATE: 2002-01-02

PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/351,814

PRIOR FILING DATE: EARLIER FILING DATE: 1999-07-12

PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: PA 1998 00937

PRIOR FILING DATE: EARLIER FILING DATE: 1998-07-15

PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: PA 1998 01667

PRIOR FILING DATE: EARLIER FILING DATE: 1998-12-17

PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/093,528

PRIOR FILING DATE: EARLIER FILING DATE: 1998-07-21

PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/115,545

PRIOR FILING DATE: EARLIER FILING DATE: 1999-01-12

NUMBER OF SEQ ID NOS: 81

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 13

LENGTH: 640

TYPE: PRT

ORGANISM: ASPERGILLUS NIGER

US-10-038-723-13

Query Match

Best Local Similarity 99.0%; Score 2742; DB 14; Length 640;

Matches 528; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy      1 MSFRLSLALSGLVCTGLANVISKRATLDSWLSNEATVARTAILNNIGADGAMVSGADSGI 60
Db      1 MSFRLSLALSGLVCTGLANVISKRATLDSWLSNEATVARTAILNNIGADGAMVSGADSGI 60
Qy      61 VVASPSNDNDPYFTYTRWRDSGLVLTLYDLFRNGDTSLSTIENYISAQAIVQGISNPSG 120
Db      61 VVASPSNDNDPYFTYTRWRDSGLVLTLYDLFRNGDTSLSTIENYISAQAIVQGISNPSG 120
Qy      121 DLSSGAGLGEPPKFNVDETAYTGSWGRPQDGPALRATAMIGFGQWLLDNGYTSATDIW 180
Db      121 DLSSGAGLGEPPKFNVDETAYTGSWGRPQDGPALRATAMIGFGQWLLDNGYTSATDIW 180
Qy      181 PLVNDLSYYAQAQWNOGTGYDLMEBVNGSSPFTTAVQHRALVEGSAFATAVGSSCSCWCDQ 240
```

```
Db      181 PLVNDLSYYAQAQWNOGTGYDLMEBVNGSSPFTTAVQHRALVEGSAFATAVGSSCSCWCDQ 240
Qy      241 APEILCYLGQFWTGSFTLIANFDSRSRSGKDNANTLLGSIHTFDPEAACDSTFQPCSPRALA 300
Db      241 APEILCYLGQFWTGSFTLIANFDSRSRSGKDNANTLLGSIHTFDPEAACDSTFQPCSPRALA 300
Qy      301 NHKEVVDSPRSIYTLNDGLSDSEAVAVGRYPEDTYNGNPMWFLCTTAAAEQLYDALYQMD 360
Db      301 NHKEVVDSPRSIYTLNDGLSDSEAVAVGRYPEDTYNGNPMWFLCTTAAAEQLYDALYQMD 360
Qy      361 KQSLFETDVSLDPEFKALYSDAATGCTYSSSSSTYSIYDAVKTPADGFVSIIVETHAASNG 420
Db      361 KQSLFETDVSLDPEFKALYSDAATGCTYSSSSSTYSIYDAVKTPADGFVSIIVETHAASNG 420
Qy      421 SMSEQYKSDGEOLASDILTWSYAALLTANNRRNSVVPASWGETSASVPCTCAATSAIG 480
Db      421 SMSEQYKSDGEOLASDILTWSYAALLTANNRRNSVVPASWGETSASVPCTCAATSAIG 480
Qy      481 TYSSTVTSWPSIYATGCTTTTATPTGSGSVTSTSKTTATATASKTSTTTRSGMSL 530
Db      481 TYSSTVTSWPSIYATGCTTTTATPTGSGSVTSTSKTTATATASKTSTTSS 530
```

RESULT 7
US-10-421-586-14

Sequence 14, Application US/10421586
Publication No. US20040002142A1

GENERAL INFORMATION:

APPLICANT: Nielsen, Bjarne Ronfeldt

APPLICANT: Svendsen, Allan

APPLICANT: Pedersen, Henrik

APPLICANT: Vind, Jesper

APPLICANT: Hendriksen, Hanne Vang

APPLICANT: Frandsen, Torben Peter

TITLE OR INVENTION: Glucoamylase Variants

FILE REFERENCE: 5967.210-US

CURRENT APPLICATION NUMBER: US/10/421,586

CURRENT FILING DATE: 2003-04-23

NUMBER OF SEQ ID NOS: 18

SOFTWARE: PatentIn version 3.2

SEQ ID NO 14

LENGTH: 640

TYPE: PRT

ORGANISM: Aspergillus niger

US-10-421-586-14

Query Match

Best Local Similarity 99.0%; Score 2742; DB 15; Length 640;

Matches 528; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy      1 MSFRLSLALSGLVCTGLANVISKRATLDSWLSNEATVARTAILNNIGADGAMVSGADSGI 60
Db      1 MSFRLSLALSGLVCTGLANVISKRATLDSWLSNEATVARTAILNNIGADGAMVSGADSGI 60
Qy      61 VVASPSNDNDPYFTYTRWRDSGLVLTLYDLFRNGDTSLSTIENYISAQAIVQGISNPSG 120
Db      61 VVASPSNDNDPYFTYTRWRDSGLVLTLYDLFRNGDTSLSTIENYISAQAIVQGISNPSG 120
Qy      121 DLSSGAGLGEPPKFNVDETAYTGSWGRPQDGPALRATAMIGFGQWLLDNGYTSATDIW 180
Db      121 DLSSGAGLGEPPKFNVDETAYTGSWGRPQDGPALRATAMIGFGQWLLDNGYTSATDIW 180
Qy      181 PLVNDLSYYAQAQWNOGTGYDLMEBVNGSSPFTTAVQHRALVEGSAFATAVGSSCSCWCDQ 240
Db      181 PLVNDLSYYAQAQWNOGTGYDLMEBVNGSSPFTTAVQHRALVEGSAFATAVGSSCSCWCDQ 240
Qy      241 APEILCYLGQFWTGSFTLIANFDSRSRSGKDNANTLLGSIHTFDPEAACDSTFQPCSPRALA 300
Db      241 APEILCYLGQFWTGSFTLIANFDSRSRSGKDNANTLLGSIHTFDPEAACDSTFQPCSPRALA 300
Qy      301 NHKEVVDSPRSIYTLNDGLSDSEAVAVGRYPEDTYNGNPMWFLCTTAAAEQLYDALYQMD 360
```

Db 301 NHKEVVDSPRSIYTLNDGLSDSEAVA VGRYPEDTYNGNPMFLCTLA AAEQLYDALYQMD 360
Qy 361 KQGSLEVTDVSLDFPKALYSDAATGTYSSSSSTYSSIVDAVKTFPADGFSIVETHAASNG 420
Db 361 KQGSLEVTDVSLDFPKALYSDAATGTYSSSSSTYSSIVDAVKTFPADGFSIVETHAASNG 420
Qy 421 SMSQYDKSDGQOLSRDLTWSYALALTANNRRNSVVPASWGETSASVPGCAATSALG 480
Db 421 SMSQYDKSDGQOLSRDLTWSYALALTANNRRNSVVPASWGETSASVPGCAATSALG 480
Qy 481 TYSSVTVTWSPIVATGTTTATPTGSGSVTSTSKTTATASKTSTTTR 530
Db 481 TYSSVTVTWSPIVATGTTTATPTGSGSVTSTSKTTATASKTSTTSS 530

RESULT 8

US-10-418-836-9
; Sequence 9, Application US/10418836
; Publication No. US20040018573A1
; GENERAL INFORMATION:
; APPLICANT: Power, Scott D.
; APPLICANT: Wang, Huaming
; APPLICANT: Ward, Michael
; TITLE OF INVENTION: Production of Functional Antibodies in
; FILE REFERENCE: GC741-2
; CURRENT APPLICATION NUMBER: US/10/418,836
; CURRENT FILING DATE: 2003-04-17
; PRIOR APPLICATION NUMBER: US 60/373,889
; PRIOR FILING DATE: 2002-04-18
; PRIOR APPLICATION NUMBER: US 60/411,540
; PRIOR FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: US 60/452,134
; PRIOR FILING DATE: 2003-03-04
; PRIOR APPLICATION NUMBER: US 60/411,537
; PRIOR FILING DATE: 2002-09-18
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 9
; LENGTH: 743
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: fusion protein
US-10-418-836-9

Query Match 98.7%; Score 2735; DB 15; Length 743;
Best Local Similarity 99.6%; Pred. No. 7.2e-234;
Matches 527; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 MSFRLSLASGLVCTGLANVISKRATLDSWLSNEATVARTAILNNIGADGAVSGADSGI 60
Db 1 MSFRLSLASGLVCTGLANVISKRATLDSWLSNEATVARTAILNNIGADGAVSGADSGI 60
Qy 61 VVASSTDNPDYFYWTMRDGLVLTLYDLFRNGDTSLSLSTIENTYISAQAIYQGISNPSG 120
Db 61 VVASSTDNPDYFYWTMRDGLVLTLYDLFRNGDTSLSLSTIENTYISAQAIYQGISNPSG 120
Qy 121 DLSGAGLGEPRKFNVDETAYTGSWGRPQRDPALATAMIGFGQWLDNGYSTATDIYW 180
Db 121 DLSGAGLGEPRKFNVDETAYTGSWGRPQRDPALATAMIGFGQWLDNGYSTATDIYW 180
Qy 181 PLVRNDLSYVAQYNNQGTGYDLMEBNVSGSFPTIAVQHRALVEGSAFATAVSGSCWCDSQ 240
Db 181 PLVRNDLSYVAQYNNQGTGYDLMEBNVSGSFPTIAVQHRALVEGSAFATAVSGSCWCDSQ 240
Qy 241 APEILCYQSFWTGSFILLANFDSRSRGKANTLLGSIHFPDEACDDSTFQPCSPRALA 300
Db 241 APEILCYQSFWTGSFILLANFDSRSRGKANTLLGSIHFPDEACDDSTFQPCSPRALA 300
Qy 301 NHKEVVDSPRSIYTLNDGLSDSEAVA VGRYPEDTYNGNPMFLCTLA AAEQLYDALYQMD 360
Db 301 NHKEVVDSPRSIYTLNDGLSDSEAVA VGRYPEDTYNGNPMFLCTLA AAEQLYDALYQMD 360

Qy 361 KQGSLEVTDVSLDFPKALYSDAATGTYSSSSSTYSSIVDAVKTFPADGFSIVETHAASNG 420
Db 361 KQGSLEVTDVSLDFPKALYSDAATGTYSSSSSTYSSIVDAVKTFPADGFSIVETHAASNG 420
Qy 421 SMSQYDKSDGQOLSRDLTWSYALALTANNRRNSVVPASWGETSASVPGCAATSALG 480
Db 421 SMSQYDKSDGQOLSRDLTWSYALALTANNRRNSVVPASWGETSASVPGCAATSALG 480
Qy 481 TYSSVTVTWSPIVATGTTTATPTGSGSVTSTSKTTATASKTSTTTR 529
Db 481 TYSSVTVTWSPIVATGTTTATPTGSGSVTSTSKTTATASKTSTYKR 529

RESULT 9

US-10-418-836-13
; Sequence 13, Application US/10418836
; Publication No. US20040018573A1
; GENERAL INFORMATION:
; APPLICANT: Power, Scott D.
; APPLICANT: Wang, Huaming
; APPLICANT: Ward, Michael
; TITLE OF INVENTION: Production of Functional Antibodies in
; FILE REFERENCE: GC741-2
; CURRENT APPLICATION NUMBER: US/10/418,836
; CURRENT FILING DATE: 2003-04-17
; PRIOR APPLICATION NUMBER: US 60/373,889
; PRIOR FILING DATE: 2002-04-18
; PRIOR APPLICATION NUMBER: US 60/411,540
; PRIOR FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: US 60/452,134
; PRIOR FILING DATE: 2003-03-04
; PRIOR APPLICATION NUMBER: US 60/411,537
; PRIOR FILING DATE: 2002-09-18
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 13
; LENGTH: 763
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: fusion protein
US-10-418-836-13

Query Match 98.7%; Score 2735; DB 15; Length 763;
Best Local Similarity 99.6%; Pred. No. 7.5e-234;
Matches 527; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 MSFRLSLASGLVCTGLANVISKRATLDSWLSNEATVARTAILNNIGADGAVSGADSGI 60
Db 1 MSFRLSLASGLVCTGLANVISKRATLDSWLSNEATVARTAILNNIGADGAVSGADSGI 60
Qy 61 VVASSTDNPDYFYWTMRDGLVLTLYDLFRNGDTSLSLSTIENTYISAQAIYQGISNPSG 120
Db 61 VVASSTDNPDYFYWTMRDGLVLTLYDLFRNGDTSLSLSTIENTYISAQAIYQGISNPSG 120
Qy 121 DLSGAGLGEPRKFNVDETAYTGSWGRPQRDPALATAMIGFGQWLDNGYSTATDIYW 180
Db 121 DLSGAGLGEPRKFNVDETAYTGSWGRPQRDPALATAMIGFGQWLDNGYSTATDIYW 180
Qy 181 PLVRNDLSYVAQYNNQGTGYDLMEBNVSGSFPTIAVQHRALVEGSAFATAVSGSCWCDSQ 240
Db 181 PLVRNDLSYVAQYNNQGTGYDLMEBNVSGSFPTIAVQHRALVEGSAFATAVSGSCWCDSQ 240
Qy 241 APEILCYQSFWTGSFILLANFDSRSRGKANTLLGSIHFPDEACDDSTFQPCSPRALA 300
Db 241 APEILCYQSFWTGSFILLANFDSRSRGKANTLLGSIHFPDEACDDSTFQPCSPRALA 300
Qy 301 NHKEVVDSPRSIYTLNDGLSDSEAVA VGRYPEDTYNGNPMFLCTLA AAEQLYDALYQMD 360
Db 301 NHKEVVDSPRSIYTLNDGLSDSEAVA VGRYPEDTYNGNPMFLCTLA AAEQLYDALYQMD 360

QY 361 KQSGLEVTDVSLDFPKALYSDAATGYSSSSSTYSIVDAVKTEADGFVSIIVETHAASNG 420
DB 361 KQSGLEVTDVSLDFPKALYSDAATGYSSSSSTYSIVDAVKTEADGFVSIIVETHAASNG 420
QY 421 SMSQYKXSGEQLSANDLTWSYAALLTANNRRNSVVPASMGETSASSVGTCAATSAIG 480
DB 421 SMSQYKXSGEQLSANDLTWSYAALLTANNRRNSVVPASMGETSASSVGTCAATSAIG 480
QY 481 TYSVTVTSMPSIVATGTTTATPTGSGSVTSTSKTTATATASKSTYTR 529
DB 481 TYSVTVTSMPSIVATGTTTATPTGSGSVTSTSKTTATATASKSTYTR 529

RESULT 10
US-10-418-836-10
; Sequence 10, Application US/10418836
; Publication No. US20040018573A1
; GENERAL INFORMATION:
; APPLICANT: Power, Scott D.
; APPLICANT: Wang, Huaming
; APPLICANT: Ward, Michael
; TITLE OF INVENTION: Production of Functional Antibodies in
; TITLE OF INVENTION: Filamentous Fungi
; FILE REFERENCE: GC741-2
; CURRENT APPLICATION NUMBER: US/10/418, 836
; CURRENT FILING DATE: 2003-04-17
; PRIOR APPLICATION NUMBER: US 60/373,889
; PRIOR FILING DATE: 2002-04-18
; PRIOR APPLICATION NUMBER: US 60/411,540
; PRIOR FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: US 60/452,134
; PRIOR FILING DATE: 2003-03-04
; PRIOR APPLICATION NUMBER: US 60/411,537
; PRIOR FILING DATE: 2002-09-18
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 979
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: fusion protein
US-10-418-836-10

Query Match 98.7%; Score 2735; DB 15; Length 979;
Best Local Similarity 99.6%; Pred. No. 1,1e-233;
Matches 527; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 MSFRSLALSGLVCTGLANVISKRATLDSWLSNEATYARATAILNNIGADGAWSGADSGI 60
DB 1 MSFRSLALSGLVCTGLANVISKRATLDSWLSNEATYARATAILNNIGADGAWSGADSGI 60
QY 61 VVASPSTNDPDPYFTWTRDSGLVLTIVDLFRNGDTSLSLTENYISAQAIVOGISNPSG 120
DB 61 VVASPSTNDPDPYFTWTRDSGLVLTIVDLFRNGDTSLSLTENYISAQAIVOGISNPSG 120
QY 121 DLSGAGLGEPRKFNVDETAYTGSWGRPQRDGPALRATAMIGFGQWLLDNGYSTATDIW 180
DB 121 DLSGAGLGEPRKFNVDETAYTGSWGRPQRDGPALRATAMIGFGQWLLDNGYSTATDIW 180
QY 181 PLYVNDLSYVAQYVNOGTGYDLMEEVNGSSFTTIAVQHRALVEGSAFATAVGSSCSWCDSQ 240
DB 181 PLYVNDLSYVAQYVNOGTGYDLMEEVNGSSFTTIAVQHRALVEGSAFATAVGSSCSWCDSQ 240
QY 241 APELICYLQSFMTGSSFTLANPDSRSGKDANTLLGSIHTFDPBAACDSTFQPCSPRALA 300
DB 241 APELICYLQSFMTGSSFTLANPDSRSGKDANTLLGSIHTFDPBAACDSTFQPCSPRALA 300
QY 301 NHKEVVDSPRSIYTLNDGLSDSEBAVAVGRYPEDITYNGNPMFLCTLLAAEQLYDALYQWD 360
DB 301 NHKEVVDSPRSIYTLNDGLSDSEBAVAVGRYPEDITYNGNPMFLCTLLAAEQLYDALYQWD 360
QY 361 KQSGLEVTDVSLDFPKALYSDAATGYSSSSSTYSIVDAVKTEADGFVSIIVETHAASNG 420

DB 361 KQSGLEVTDVSLDFPKALYSDAATGYSSSSSTYSIVDAVKTEADGFVSIIVETHAASNG 420
QY 421 SMSQYKXSGEQLSANDLTWSYAALLTANNRRNSVVPASMGETSASSVGTCAATSAIG 480
DB 421 SMSQYKXSGEQLSANDLTWSYAALLTANNRRNSVVPASMGETSASSVGTCAATSAIG 480
QY 481 TYSVTVTSMPSIVATGTTTATPTGSGSVTSTSKTTATATASKSTYTR 529
DB 481 TYSVTVTSMPSIVATGTTTATPTGSGSVTSTSKTTATATASKSTYTR 529

RESULT 11
US-10-418-836-16
; Sequence 16, Application US/10418836
; Publication No. US20040018573A1
; GENERAL INFORMATION:
; APPLICANT: Power, Scott D.
; APPLICANT: Wang, Huaming
; APPLICANT: Ward, Michael
; TITLE OF INVENTION: Production of Functional Antibodies in
; TITLE OF INVENTION: Filamentous Fungi
; FILE REFERENCE: GC741-2
; CURRENT APPLICATION NUMBER: US/10/418, 836
; CURRENT FILING DATE: 2003-04-17
; PRIOR APPLICATION NUMBER: US 60/373,889
; PRIOR FILING DATE: 2002-04-18
; PRIOR APPLICATION NUMBER: US 60/411,540
; PRIOR FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: US 60/452,134
; PRIOR FILING DATE: 2003-03-04
; PRIOR APPLICATION NUMBER: US 60/411,537
; PRIOR FILING DATE: 2002-09-18
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 979
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: fusion protein
US-10-418-836-16

Query Match 98.7%; Score 2735; DB 15; Length 979;
Best Local Similarity 99.6%; Pred. No. 1,1e-233;
Matches 527; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 MSFRSLALSGLVCTGLANVISKRATLDSWLSNEATYARATAILNNIGADGAWSGADSGI 60
DB 1 MSFRSLALSGLVCTGLANVISKRATLDSWLSNEATYARATAILNNIGADGAWSGADSGI 60
QY 61 VVASPSTNDPDPYFTWTRDSGLVLTIVDLFRNGDTSLSLTENYISAQAIVOGISNPSG 120
DB 61 VVASPSTNDPDPYFTWTRDSGLVLTIVDLFRNGDTSLSLTENYISAQAIVOGISNPSG 120
QY 121 DLSGAGLGEPRKFNVDETAYTGSWGRPQRDGPALRATAMIGFGQWLLDNGYSTATDIW 180
DB 121 DLSGAGLGEPRKFNVDETAYTGSWGRPQRDGPALRATAMIGFGQWLLDNGYSTATDIW 180
QY 181 PLYVNDLSYVAQYVNOGTGYDLMEEVNGSSFTTIAVQHRALVEGSAFATAVGSSCSWCDSQ 240
DB 181 PLYVNDLSYVAQYVNOGTGYDLMEEVNGSSFTTIAVQHRALVEGSAFATAVGSSCSWCDSQ 240
QY 241 APELICYLQSFMTGSSFTLANPDSRSGKDANTLLGSIHTFDPBAACDSTFQPCSPRALA 300
DB 241 APELICYLQSFMTGSSFTLANPDSRSGKDANTLLGSIHTFDPBAACDSTFQPCSPRALA 300
QY 301 NHKEVVDSPRSIYTLNDGLSDSEBAVAVGRYPEDITYNGNPMFLCTLLAAEQLYDALYQWD 360
DB 301 NHKEVVDSPRSIYTLNDGLSDSEBAVAVGRYPEDITYNGNPMFLCTLLAAEQLYDALYQWD 360
QY 361 KQSGLEVTDVSLDFPKALYSDAATGYSSSSSTYSIVDAVKTEADGFVSIIVETHAASNG 420

Db 361 KQGSLEVTVDLSLDFPKALYSDAATGYSSSSSTYSIVDAVKTADGFSIVETHAASNG 420
 QY 421 SMSSEYDSDGSEQLSARDLTWSYAAALLTANNRRNSVVPASWGETSASSVPGCAATSAIG 480
 Db 421 SMSSEYDSDGSEQLSARDLTWSYAAALLTANNRRNSVVPASWGETSASSVPGCAATSAIG 480
 QY 481 TVSSVTVTWSPIVATGTTTATPTGSGSVTSTSKTTATASKTSTTR 529
 Db 481 TVSSVTVTWSPIVATGTTTATPTGSGSVTSTSKTTATASKTSTTR 529

RESULT 12

US-10-418-836-19
 ; Sequence 19, Application US/10418836
 ; Publication No. US20040018573A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Power, Scott D.
 ; APPLICANT: Wang, Huaming
 ; APPLICANT: Ward, Michael
 ; TITLE OF INVENTION: Production of Functional Antibodies in
 ; TITLE OF INVENTION: Filamentous Fungi
 ; FILE REFERENCE: GC741-2
 ; CURRENT APPLICATION NUMBER: US/10/418,836
 ; CURRENT FILING DATE: 2003-04-17
 ; PRIOR APPLICATION NUMBER: US 60/373,889
 ; PRIOR FILING DATE: 2002-04-18
 ; PRIOR APPLICATION NUMBER: US 60/411,540
 ; PRIOR FILING DATE: 2002-09-18
 ; PRIOR APPLICATION NUMBER: US 60/452,134
 ; PRIOR FILING DATE: 2003-03-04
 ; PRIOR APPLICATION NUMBER: US 60/411,537
 ; PRIOR FILING DATE: 2002-09-18
 ; NUMBER OF SEQ ID NOS: 40
 ; SOFTWARE: FaestSeq for Windows Version 4.0
 ; SEQ ID NO 19
 ; LENGTH: 738
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: fusion protein
 ; US-10-418-836-19

Query Match 98.2%; Score 2720.5; DB 15; Length 738;
 Best Local Similarity 98.5%; Pred. No. 1.4e-232;
 Matches 527; Conservative 1; Mismatches 6; Indels 1; Gaps 1;

QY 1 MSFRLSLALSGLVCTGLANVISKRATLDSWLSNEATVARTAILNNIGADGAVVSGADSGI 60
 Db 1 MSFRLSLALSGLVCTGLANVISKRATLDSWLSNEATVARTAILNNIGADGAVVSGADSGI 60
 QY 61 VVASPTSTNPDPFYTWTRDSGLVLTLYDLFRNGDTSLSLTIENTISAQAIYQGISNPSG 120
 Db 61 VVASPTSTNPDPFYTWTRDSGLVLTLYDLFRNGDTSLSLTIENTISAQAIYQGISNPSG 120
 QY 121 DLSGAGLGEPEKFNVDERTAYTSGMGRPORDGPALRATAMIGFGWLLONGYTSRTDLYW 180
 Db 121 DLSGAGLGEPEKFNVDERTAYTSGMGRPORDGPALRATAMIGFGWLLONGYTSRTDLYW 180
 QY 121 DLSGAGLGEPEKFNVDERTAYTSGMGRPORDGPALRATAMIGFGWLLONGYTSRTDLYW 180
 Db 121 DLSGAGLGEPEKFNVDERTAYTSGMGRPORDGPALRATAMIGFGWLLONGYTSRTDLYW 180
 QY 181 PLVRNDLSYVAQYNNQGYDLMEEVNGSSFTTIAVOHRALVEGSAFATAVSSSCSWCDSQ 240
 Db 181 PLVRNDLSYVAQYNNQGYDLMEEVNGSSFTTIAVOHRALVEGSAFATAVSSSCSWCDSQ 240
 QY 241 APELICYSQFWTGSFILANPDSRSKGDANTLLGSIHTFDEAACDSTFQPCSPRALA 300
 Db 241 APELICYSQFWTGSFILANPDSRSKGDANTLLGSIHTFDEAACDSTFQPCSPRALA 300
 QY 301 NHKEVVDSEFRSITYTLNDGLSDSEAVAVGRYPEDTYNNGNPMFLCTLLAAEQLYDALYQMD 360
 Db 301 NHKEVVDSEFRSITYTLNDGLSDSEAVAVGRYPEDTYNNGNPMFLCTLLAAEQLYDALYQMD 360
 QY 361 KQGSLEVTVDLSLDFPKALYSDAATGYSSSSSTYSIVDAVKTADGFSIVETHAASNG 420
 Db 361 KQGSLEVTVDLSLDFPKALYSDAATGYSSSSSTYSIVDAVKTADGFSIVETHAASNG 420

QY 421 SMSSEYDSDGSEQLSARDLTWSYAAALLTANNRRNSVVPASWGETSASSVPGCAATSAIG 480
 Db 421 SMSSEYDSDGSEQLSARDLTWSYAAALLTANNRRNSVVPASWGETSASSVPGCAATSAIG 480
 QY 481 TVSSVTVTWSPIVATGTTTATPTGSGSVTSTSKTTATASKTSTTR 534
 Db 481 TVSSVTVTWSPIVATGTTTATPTGSGSVTSTSKTTATASKTSTTR 534

RESULT 13

US-10-418-836-30
 ; Sequence 30, Application US/10418836
 ; Publication No. US20040018573A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Power, Scott D.
 ; APPLICANT: Wang, Huaming
 ; APPLICANT: Ward, Michael
 ; TITLE OF INVENTION: Production of Functional Antibodies in
 ; TITLE OF INVENTION: Filamentous Fungi
 ; FILE REFERENCE: GC741-2
 ; CURRENT APPLICATION NUMBER: US/10/418,836
 ; CURRENT FILING DATE: 2003-04-17
 ; PRIOR APPLICATION NUMBER: US 60/373,889
 ; PRIOR FILING DATE: 2002-04-18
 ; PRIOR APPLICATION NUMBER: US 60/411,540
 ; PRIOR FILING DATE: 2002-09-18
 ; PRIOR APPLICATION NUMBER: US 60/452,134
 ; PRIOR FILING DATE: 2003-03-04
 ; PRIOR APPLICATION NUMBER: US 60/411,537
 ; PRIOR FILING DATE: 2002-09-18
 ; NUMBER OF SEQ ID NOS: 40
 ; SOFTWARE: FaestSeq for Windows Version 4.0
 ; SEQ ID NO 30
 ; LENGTH: 741
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: fusion protein
 ; US-10-418-836-30

Query Match 98.0%; Score 2716; DB 15; Length 741;
 Best Local Similarity 98.0%; Pred. No. 3.5e-232;
 Matches 527; Conservative 1; Mismatches 6; Indels 4; Gaps 1;

QY 1 MSFRLSLALSGLVCTGLANVISKRATLDSWLSNEATVARTAILNNIGADGAVVSGADSGI 60
 Db 1 MSFRLSLALSGLVCTGLANVISKRATLDSWLSNEATVARTAILNNIGADGAVVSGADSGI 60
 QY 61 VVASPTSTNPDPFYTWTRDSGLVLTLYDLFRNGDTSLSLTIENTISAQAIYQGISNPSG 120
 Db 61 VVASPTSTNPDPFYTWTRDSGLVLTLYDLFRNGDTSLSLTIENTISAQAIYQGISNPSG 120
 QY 121 DLSGAGLGEPEKFNVDERTAYTSGMGRPORDGPALRATAMIGFGWLLONGYTSRTDLYW 180
 Db 121 DLSGAGLGEPEKFNVDERTAYTSGMGRPORDGPALRATAMIGFGWLLONGYTSRTDLYW 180
 QY 121 DLSGAGLGEPEKFNVDERTAYTSGMGRPORDGPALRATAMIGFGWLLONGYTSRTDLYW 180
 Db 121 DLSGAGLGEPEKFNVDERTAYTSGMGRPORDGPALRATAMIGFGWLLONGYTSRTDLYW 180
 QY 181 PLVRNDLSYVAQYNNQGYDLMEEVNGSSFTTIAVOHRALVEGSAFATAVSSSCSWCDSQ 240
 Db 181 PLVRNDLSYVAQYNNQGYDLMEEVNGSSFTTIAVOHRALVEGSAFATAVSSSCSWCDSQ 240
 QY 241 APELICYSQFWTGSFILANPDSRSKGDANTLLGSIHTFDEAACDSTFQPCSPRALA 300
 Db 241 APELICYSQFWTGSFILANPDSRSKGDANTLLGSIHTFDEAACDSTFQPCSPRALA 300
 QY 301 NHKEVVDSEFRSITYTLNDGLSDSEAVAVGRYPEDTYNNGNPMFLCTLLAAEQLYDALYQMD 360
 Db 301 NHKEVVDSEFRSITYTLNDGLSDSEAVAVGRYPEDTYNNGNPMFLCTLLAAEQLYDALYQMD 360
 QY 361 KQGSLEVTVDLSLDFPKALYSDAATGYSSSSSTYSIVDAVKTADGFSIVETHAASNG 420
 Db 361 KQGSLEVTVDLSLDFPKALYSDAATGYSSSSSTYSIVDAVKTADGFSIVETHAASNG 420

[illegible]

RESULT 14
US-10-418

```

Sequence 38, Application US/10418836
Publication No. US20040018573A1
GENERAL INFORMATION:
APPLICANT: Power, Scott D.
APPLICANT: Wang, Huang
APPLICANT: Ward, Michael
TITLE OF INVENTION: Production of Fungitio
TITLE OF INVENTION: Filamentous Fungi
FILE REFERENCE: GC/41-2
CURRENT APPLICATION NUMBER: US/10/418, 836
CURRENT FILING DATE: 2003-04-17
PRIOR APPLICATION NUMBER: US 60/373, 889
PRIOR FILING DATE: 2002-04-18
PRIOR APPLICATION NUMBER: US 60/411, 540
PRIOR FILING DATE: 2002-09-18
PRIOR APPLICATION NUMBER: US 60/452, 134
PRIOR FILING DATE: 2003-03-04
PRIOR APPLICATION NUMBER: US 60/411, 537
PRIOR FILING DATE: 2002-09-18
NUMBER OF SEQ ID NOS: 40
SOFTWARE: Fastseq for Windows Version 4.0
SEQ ID NO 38
LENGTH: 972
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: fusion protein
US-10-418-836-38

```

Query Match	Score	DB	Length
98.0%	2715.5	15	972

Best Local Similarity 95.5%; Pred.No. 5,8e-232;
Matches 530; Conservative 0; Mismatches 4; Indels 21; Gaps 2;

Qy	1	MSRSLIALSGVLCGTGLANVISKRAITLDSWLSNEATVARTALLNNIGADGAWGADSGI	60
Dd	1	MSRSLIALSGVLCGTGLANVISKRAITLDSWLSNEATVARTALLNNIGADGAWGADSGI	60
Qy	61	VVASPSTDNDEYFYTTWTRDSGLVLTLYDLFPNKGDTSLSTIENYISQAIVQGISNPSG	120
Dd	61	VVASPSTDNDEYFYTTWTRDSGLVLTLYDLFPNKGDTSLSTIENYISQAIVQGISNPSG	120
Qy	121	DISSGAGLGEPKFNVDETAVTGSGWGRPQKDGALATATMIGFGQWLDNGYTATDIYW	180
Dd	121	DISSGAGLGEPKFNVDETAVTGSGWGRPQKDGALATATMIGFGQWLDNGYTATDIYW	180
Qy	181	PLVRLNDLSYAAQVYNNQGYDLMEEVNGSSFPIIANQHALVEGSAFATAVSSCSWCDSQ	240
Dd	181	PLVRLNDLSYAAQVYNNQGYDLMEEVNGSSFPIIANQHALVEGSAFATAVSSCSWCDSQ	240
Qy	241	APELLICLOSFWMGSCFILANPDSRSRGKANTLLGSIHTPPEBAACDSTPQCSPALA	300
Dd	241	APELLICLOSFWMGSCFILANPDSRSRGKANTLLGSIHTPPEBAACDSTPQCSPALA	300
Qy	301	NHKEVVDSPRSIYTLNDGLSDSEAAVAVGRYPEDTYNNGNPMFLCTIAAEOUYALQYMD	360
Dd	301	NHKEVVDSPRSIYTLNDGLSDSEAAVAVGRYPEDTYNNGNPMFLCTIAAEOUYALQYMD	360
Qy	361	KQGSLEVTIVDSLDPFALYSDAATGYSSSSSTYSGSIYDAVYTFADGVSIIVETHAASNG	420
Dd	361	KQGSLEVTIVDSLDPFALYSDAATGYSSSSSTYSGSIYDAVYTFADGVSIIVETHAASNG	420
Qy	421	SMSBOYDKSDGEQLASRDITWSYAAALLTNNRNSVVPASMGETSASSVPQTCATSAIG	480

Db 421 SMEQYDKSDGQLSARDLTWSEYALLLRNNRRNSVYPASGGEISASVSQTCAATSAIG 480
 Qy 481 TYSSTVTWSPBSIVATGGTTTATPTGGGASTTSKTTATKS-----KT 524
 Db 481 TYSSTVTWSPBSIVATGGTTTATPTGGGASTTSKTTATKSISKROVOLAQSGPGLVKP 540
 Qy 525 S-----TTRSGMSL 534
 Db 541 SEFTLSLTCTVSGPGL 555

RESULT 15

```

US-10-418-836-26
? Sequence 26, Application US/10418636
? Publication No. US20040018573A1
? GENERAL INFORMATION:
? APPLICANT: Power, Scott D.
? APPLICANT: Wang, Huaming
? APPLICANT: Ward, Michael
? TITLE OF INVENTION: Production of Function
? TITLE OF INVENTION: Filamentous Fungi
? FILE REFERENCE: GCF41-2
? CURRENT APPLICATION NUMBER: US/10/418, 836
? CURRENT FILING DATE: 2003-04-17
? PRIOR APPLICATION NUMBER: US 60/373, 889
? PRIOR FILING DATE: 2002-04-18
? PRIOR APPLICATION NUMBER: US 60/411, 540
? PRIOR FILING DATE: 2002-09-18
? PRIOR APPLICATION NUMBER: US 60/452, 134
? PRIOR FILING DATE: 2003-03-04
? PRIOR APPLICATION NUMBER: US 60/411, 537
? PRIOR FILING DATE: 2002-09-18
? NUMBER OF SEQ ID NOS: 40
? SOFTWARE: RebaseSeq for Windows Version 4.0.
? SEQ ID NO 26
? LENGTH: 739
? TYPE: PRT
? ORGANISM: Artificial Sequence
? FEATURE:
? OTHER INFORMATION: fusion protein
US-10-418-836-26

```

Query Match	98.0%	Score 2715	DB 15	Length 739
-------------	-------	------------	-------	------------

Best Local Similarity 99.8%; Pred. No. 4,3e-232;
Matches 522; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MSRPSLLATSLG.VCTGLANVISKRATLDSML.SNEATVARTAILNNICADGAMVSGADSGI	60
Dp	1	MSRPSLLATSLG.VCTGLANVISKRATLDSML.SNEATVARTAILNNICADGAMVSGADSGI	60
Qy	61	VVASPSTNDPDPFYTWTRDSGLVLTLLVDLFRNGDTSILSTIENYISAQAI.VOGISNPSG	120
Dp	61	VVASPSTNDPDPFYTWTRDSGLVLTLLVDLFRNGDTSILSTIENYISAQAI.VOGISNPSG	120
Qy	121	DLSSGAGLCEPKFNDEIATYGSWGR.PORDGALPALTAMIGFGWLLDNGYSTATDIYW	180
Dp	121	DLSSGAGLCEPKFNDEIATYGSWGR.PORDGALPALTAMIGFGWLLDNGYSTATDIYW	180
Qy	181	PLVRNDLSVAAQYMNQGYDLMEEEVNGSSPFILIVQHRALVEGSAFATVAGSSCSWCDSQ	240
Dp	181	PLVRNDLSVAAQYMNQGYDLMEEEVNGSSPFILIVQHRALVEGSAFATVAGSSCSWCDSQ	240
Qy	241	APRILCYLOSFWTGSFILLANFPDSSRSRGDANTLLGSIHTPDEPBAACDSTFPOCSPRALA	300
Dp	241	APRILCYLOSFWTGSFILLANFPDSSRSRGDANTLLGSIHTPDEPBAACDSTFPOCSPRALA	300
Qy	301	NHKEVVDSPRSIYTLINDGISDSEAAVAVGRYPEDTYNGNPNWFLCTLLAAEQLYDALYQWD	360
Dp	301	NHKEVVDSPRSIYTLINDGISDSEAAVAVGRYPEDTYNGNPNWFLCTLLAAEQLYDALYQWD	360
Qy	361	KQGSLEVDVSLDFFKALYSDAATGTYSSSSSTYSIYDANKVTRADGQVSI.VETHAASNG	420
Dp	361	KQGSLEVDVSLDFFKALYSDAATGTYSSSSSTYSIYDANKVTRADGQVSI.VETHAASNG	420

Db	361	KQGSLEVTDVSLDPFKALYSDAATGYSSSSSTYSIVDAVKTADGFVSIETHASNG	420
Qy	421	SMSEQYDKSDGQOLASARDLTWSYALLTANNRRNSVVPASWGETGASSVPGTCAATSAG	480
Db	421	SMSEQYDKSDGQOLASARDLTWSYALLTANNRRNSVVPASWGETSASSVPGTCAATSAG	480
Qy	481	TYSSVTVTSMPSIVATGGTTTATPTGSGSVTSTSKTTATASK	523
Db	481	TYSSVTVTSMPSIVATGGTTTATPTGSGSVTSTSKTTATASQ	523

Search completed: June 17, 2004, 17:53:29
 Job time : 51 secs

2110 PAGE LEFT BLANK

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: June 17, 2004, 17:44:49 ; Search time 23 Seconds
(without alignments)
1198.620 Million cell updates/sec

Title: US-10-038-723-2
Perfect score: 2771
Sequence: 1 MSFRLSLALSLGLVCTGLANV.....SKTTATASKSTTTRSGMSL 534

Scoring table: BLASTSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/6C_COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2771	100.0	534	US-09-198-672-2	Sequence 2, Appl1
2	2771	100.0	534	US-09-199-290-9	Sequence 9, Appl1
3	2771	100.0	534	US-09-632-392-2	Sequence 2, Appl1
4	2771	100.0	534	US-09-455-679-1	Sequence 1, Appl1
5	2771	100.0	534	US-09-351-814-2	Sequence 2, Appl1
6	2771	100.0	534	US-09-821-616-9	Sequence 9, Appl1
7	2742	99.0	640	US-09-351-814-13	Sequence 13, Appl1
8	2628	94.8	616	US-09-236-063-1	Sequence 1, Appl1
9	1689	61.0	618	US-09-199-290-34	Sequence 34, Appl1
10	1689	61.0	618	US-09-821-616-34	Sequence 34, Appl1
11	1674.5	60.4	591	US-09-199-290-7	Sequence 7, Appl1
12	1674.5	60.4	591	US-09-821-616-7	Sequence 7, Appl1
13	1615.5	58.3	630	US-09-704-449-2	Sequence 2, Appl1
14	1577	56.9	626	US-08-596-300A-7	Sequence 7, Appl1
15	1577	56.9	626	US-08-596-300A-14	Sequence 14, Appl1
16	1425.5	51.4	581	US-09-534-407-3	Sequence 3, Appl1
17	1425.5	51.4	581	US-09-999-201B-4	Sequence 4, Appl1
18	1321	47.7	616	US-08-385-370-2	Sequence 2, Appl1
19	1321	47.7	616	US-08-385-370-4	Sequence 4, Appl1
20	589	21.3	806	US-08-270-076A-11	Sequence 11, Appl1
21	160	5.8	35	US-09-199-290-4	Sequence 4, Appl1
22	160	5.8	35	US-09-821-616-4	Sequence 4, Appl1
23	156.5	5.6	2137	US-09-134-001C-4463	Sequence 4463, Ap
24	142	5.1	536	US-09-463-712C-10	Sequence 10, Appl1
25	136	4.9	1026	US-08-614-377A-7	Sequence 7, Appl1
26	136	4.9	1026	US-09-142-648B-7	Sequence 7, Appl1
27	133	4.8	1529	US-08-728-470-10	Sequence 10, Appl1

28	133	4.8	1529	3	US-08-719-641-10	Sequence 10, Appl1
29	132	4.8	1426	3	US-09-136-574A-43	Sequence 43, Appl1
30	130.5	4.7	1537	1	US-08-325-267A-2	Sequence 2, Appl1
31	130	4.7	1600	2	US-08-617-687-10	Sequence 10, Appl1
32	129	4.7	1026	2	US-08-194-290-7	Sequence 7, Appl1
33	128.5	4.6	1721	3	US-09-107-532A-5134	Sequence 5134, Ap
34	128	4.6	1721	3	US-08-700-651-5	Sequence 5, Appl1
35	128	4.6	1721	3	US-08-928-361B-6	Sequence 6, Appl1
36	128	4.6	1721	4	US-09-588-995A-6	Sequence 6, Appl1
37	127.5	4.6	751	6	5457037-5	Patent No. 5457037
38	127.5	4.6	776	3	US-09-346-237-7	Sequence 7, Appl1
39	127	4.6	630	2	US-08-797-366-3	Sequence 3, Appl1
40	127	4.6	630	2	US-08-956-268-3	Sequence 3, Appl1
41	127	4.6	906	4	US-09-717-364A-15	Sequence 15, Appl1
42	126.5	4.6	776	3	US-09-346-237-4	Sequence 4, Appl1
43	126	4.5	288	4	US-09-216-393B-341	Sequence 341, App
44	126	4.5	288	4	US-09-216-393B-344	Sequence 344, App
45	124.5	4.5	894	3	US-08-362-525-22	Sequence 22, Appl1

ALIGNMENTS

```
RESULT 1
US-09-198-672-2
; Sequence 2, Application US/09198672
; Patent No. 6129788
; GENERAL INFORMATION:
; APPLICANT: Liaw, Gin
; APPLICANT: Pedersen, Sven
; TITLE OF INVENTION: A Method of Producing Saccharide
; TITLE OF INVENTION: Preparations
; FILE REFERENCE: 5318,200-US
; CURRENT APPLICATION NUMBER: US/09/198, 672
; CURRENT FILING DATE: 1998-11-23
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 534
; TYPE: PRT
; ORGANISM: Aspergillus Niger
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)...(24)
US-09-198-672-2
Query Match          100.0%; Score 2771; DB 3; Length 534;
Best Local Similarity 100.0%; Pred. No. 1.1e-211;
Matches 534; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 MSFRLSLALSLGLVCTGLANVISKRATLDSWLSNEATVARTAILNNIGADGAVSGADSGI 60
DB      1 MSFRLSLALSLGLVCTGLANVISKRATLDSWLSNEATVARTAILNNIGADGAVSGADSGI 60
QY      61 VVASPSDNPYFYTWTRDGLVLTVDLPRNDTSLISTEYISQAIVQGISNSG 120
DB      61 VVASPSDNPYFYTWTRDGLVLTVDLPRNDTSLISTEYISQAIVQGISNSG 120
QY      121 DLSSGAGIGERKFNVDATAVYSGMGRPORDPALRATAMIFGQWLDLNGYSTATDVIW 180
DB      121 DLSSGAGIGERKFNVDATAVYSGMGRPORDPALRATAMIFGQWLDLNGYSTATDVIW 180
QY      181 PLVNDLSYVAQYNQGYDIMEEVNGSSFFTIAVQHRALVEGSAFATVAGSSCSWCDSQ 240
DB      181 PLVNDLSYVAQYNQGYDIMEEVNGSSFFTIAVQHRALVEGSAFATVAGSSCSWCDSQ 240
QY      241 APELICVQSWTQSFILANDSSRGKNDATLLGSIHTPDPEAACDSTFQPCSPRLA 300
DB      241 APELICVQSWTQSFILANDSSRGKNDATLLGSIHTPDPEAACDSTFQPCSPRLA 300
QY      301 NHRKVVNSFRSITYTLNDGLDSEBAVAVGRVPEDTYNGNPFILCTLLAAEQLYALYQWD 360
DB      301 NHRKVVNSFRSITYTLNDGLDSEBAVAVGRVPEDTYNGNPFILCTLLAAEQLYALYQWD 360
```

```
Db 301 NHKEVDSFRSITYLTINDGLSDSEAVAVGRYPEDTYNGNPMWFLCTTAAAEQLYDALYQMD 360
Qy 361 KQGSLEVTDVSLDFPKALYSDAATGTYSSTSSSTYSIVDAVKTFADGFVSIIVETHAASNG 420
Db 361 KQGSLEVTDVSLDFPKALYSDAATGTYSSTSSSTYSIVDAVKTFADGFVSIIVETHAASNG 420
Qy 421 SMSBOYKSDGEQUSARDLTWVSYAALLTANNRRNSVVPASMGETSASVPGTCAATSAIG 480
Db 421 SMSBOYKSDGEQUSARDLTWVSYAALLTANNRRNSVVPASMGETSASVPGTCAATSAIG 480
Qy 481 TYSVSVTWSMPSIVATGTTTATPTGSGSVTSTSKTTATASKTSTTTRSGMSL 534
Db 481 TYSVSVTWSMPSIVATGTTTATPTGSGSVTSTSKTTATASKTSTTTRSGMSL 534

RESULT 2
US-09-199-290-9
; Sequence 9, Application US/09199290
; Patent No. 6255084
; GENERAL INFORMATION:
; APPLICANT: Nielsen, Bjarne R.
; APPLICANT: Nielsen, Rudy
; APPLICANT: Lehmebeck, Jan
; TITLE OF INVENTION: Thermostable Glucoamylase
; FILE REFERENCE: 5279, 200-US
; CURRENT APPLICATION NUMBER: US/09/199, 290
; EARLIER FILING DATE: 1998-11-24
; EARLIER APPLICATION NUMBER: 1557/97
; EARLIER FILING DATE: 1997-12-30
; EARLIER APPLICATION NUMBER: 0925/98
; EARLIER FILING DATE: 1998-07-10
; EARLIER APPLICATION NUMBER: 60/070, 746
; EARLIER FILING DATE: 1998-01-08
; EARLIER APPLICATION NUMBER: 60/094, 344
; EARLIER FILING DATE: 1998-07-28
; EARLIER APPLICATION NUMBER: 08/979, 673
; EARLIER FILING DATE: 1997-11-26
; EARLIER APPLICATION NUMBER: 09/107, 657
; EARLIER FILING DATE: 1998-06-30
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 9
; LENGTH: 534
; TYPE: PRT
; ORGANISM: Aspergillus niger
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)...(24)
US-09-199-290-9

Query Match 100.0%; Score 2771; DB 3; Length 534;
Best Local Similarity 100.0%; Pred. No. 1,1e-231;
Matches 534; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSFRLSLALSGLVCTGLANVISKRATLDSWLSNEATVARTAILNNIGADGAWSGADSGI 60
Db 1 MSFRLSLALSGLVCTGLANVISKRATLDSWLSNEATVARTAILNNIGADGAWSGADSGI 60
Qy 61 VVASPSTNDNDYFTYTWTRDSGLVYKTLVLDLFRNGDTSLSTIENTYISAQAIVOGISNPSG 120
Db 61 VVASPSTNDNDYFTYTWTRDSGLVYKTLVLDLFRNGDTSLSTIENTYISAQAIVOGISNPSG 120
Qy 121 DLSSGAGLGEPRKFVNDETATYTGSMGRPORDBPALRATAMIGFGOMLNDNGTSTATDIW 180
Db 121 DLSSGAGLGEPRKFVNDETATYTGSMGRPORDBPALRATAMIGFGOMLNDNGTSTATDIW 180
Qy 181 PLVNDLSYVAQYVNGQGYDLMEEVNGSSFTTIAVQHRALVEGSAFATAVSSCSWCDSQ 240
Db 181 PLVNDLSYVAQYVNGQGYDLMEEVNGSSFTTIAVQHRALVEGSAFATAVSSCSWCDSQ 240
Qy 241 APELICLYQSFWTGSLFIANFDSRSRSGKDANTLLGSIHTFDPEAACDSTFQPCSPRALA 300
Db 241 APELICLYQSFWTGSLFIANFDSRSRSGKDANTLLGSIHTFDPEAACDSTFQPCSPRALA 300
```

```
Qy 301 NHKEVDSFRSITYLTINDGLSDSEAVAVGRYPEDTYNGNPMWFLCTTAAAEQLYDALYQMD 360
Db 301 NHKEVDSFRSITYLTINDGLSDSEAVAVGRYPEDTYNGNPMWFLCTTAAAEQLYDALYQMD 360
Qy 361 KQGSLEVTDVSLDFPKALYSDAATGTYSSTSSSTYSIVDAVKTFADGFVSIIVETHAASNG 420
Db 361 KQGSLEVTDVSLDFPKALYSDAATGTYSSTSSSTYSIVDAVKTFADGFVSIIVETHAASNG 420
Qy 421 SMSBOYKSDGEQUSARDLTWVSYAALLTANNRRNSVVPASMGETSASVPGTCAATSAIG 480
Db 421 SMSBOYKSDGEQUSARDLTWVSYAALLTANNRRNSVVPASMGETSASVPGTCAATSAIG 480
Qy 481 TYSVSVTWSMPSIVATGTTTATPTGSGSVTSTSKTTATASKTSTTTRSGMSL 534
Db 481 TYSVSVTWSMPSIVATGTTTATPTGSGSVTSTSKTTATASKTSTTTRSGMSL 534

RESULT 3
US-09-632-392-2
; Sequence 2, Application US/09632392
; Patent No. 6303346
; GENERAL INFORMATION:
; APPLICANT: Liaw, Gin
; APPLICANT: Pedersen, Sven
; APPLICANT: Hendriksen, Sven
; TITLE OF INVENTION: A Method of Producing Saccharide
; TITLE OF INVENTION: Preparations
; FILE REFERENCE: 5318, 200-US
; CURRENT APPLICATION NUMBER: US/09/632, 392
; EARLIER FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 09/198, 672
; PRIOR FILING DATE: 1998-11-23
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 534
; TYPE: PRT
; ORGANISM: Aspergillus Niger
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)...(24)
US-09-632-392-2

Query Match 100.0%; Score 2771; DB 4; Length 534;
Best Local Similarity 100.0%; Pred. No. 1,1e-231;
Matches 534; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSFRLSLALSGLVCTGLANVISKRATLDSWLSNEATVARTAILNNIGADGAWSGADSGI 60
Db 1 MSFRLSLALSGLVCTGLANVISKRATLDSWLSNEATVARTAILNNIGADGAWSGADSGI 60
Qy 61 VVASPSTNDNDYFTYTWTRDSGLVYKTLVLDLFRNGDTSLSTIENTYISAQAIVOGISNPSG 120
Db 61 VVASPSTNDNDYFTYTWTRDSGLVYKTLVLDLFRNGDTSLSTIENTYISAQAIVOGISNPSG 120
Qy 121 DLSSGAGLGEPRKFVNDETATYTGSMGRPORDBPALRATAMIGFGOMLNDNGTSTATDIW 180
Db 121 DLSSGAGLGEPRKFVNDETATYTGSMGRPORDBPALRATAMIGFGOMLNDNGTSTATDIW 180
Qy 181 PLVNDLSYVAQYVNGQGYDLMEEVNGSSFTTIAVQHRALVEGSAFATAVSSCSWCDSQ 240
Db 181 PLVNDLSYVAQYVNGQGYDLMEEVNGSSFTTIAVQHRALVEGSAFATAVSSCSWCDSQ 240
Qy 241 APELICLYQSFWTGSLFIANFDSRSRSGKDANTLLGSIHTFDPEAACDSTFQPCSPRALA 300
Db 241 APELICLYQSFWTGSLFIANFDSRSRSGKDANTLLGSIHTFDPEAACDSTFQPCSPRALA 300
Qy 301 NHKEVDSFRSITYLTINDGLSDSEAVAVGRYPEDTYNGNPMWFLCTTAAAEQLYDALYQMD 360
Db 301 NHKEVDSFRSITYLTINDGLSDSEAVAVGRYPEDTYNGNPMWFLCTTAAAEQLYDALYQMD 360
Qy 361 KQGSLEVTDVSLDFPKALYSDAATGTYSSTSSSTYSIVDAVKTFADGFVSIIVETHAASNG 420
```

Db 361 KQGSLEVTVDLDFPKALYSDPATGTYSSTSSSTYSIVDAVKTFPDGFVSIYETHAASNG 420
Qy 421 SMSSEQYDKSDGEQSLASRLDTSWYALLLTANNRRNSVVPASWGETSASSVPGTCAATSAIG 480
Db 421 SMSSEQYDKSDGEQSLASRLDTSWYALLLTANNRRNSVVPASWGETSASSVPGTCAATSAIG 480
Qy 481 TYSSTVTVSWPSIYAVGTGTTTATPTGSGSVTSKTKTATASKSTTTTRSGMSL 534
Db 481 TYSSTVTVSWPSIYAVGTGTTTATPTGSGSVTSKTKTATASKSTTTTRSGMSL 534

RESULT 4
US-09-455-679-1
Sequence 1, Application US/09455679

Patent No. 6329186
GENERAL INFORMATION:
APPLICANT: Nielsen, Bjarne R.
APPLICANT: Svendsen, Allan
APPLICANT: Bojlsen, Kirsten
APPLICANT: Vind, Jesper
APPLICANT: Pedersen, Henrik
TITLE OF INVENTION: Glucoamyase with N-Terminal Extensions
FILE REFERENCE: 5691.200-US
CURRENT FILING DATE: 1999-12-07
EARLIER FILING DATE: 1998-12-07
EARLIER FILING DATE: 1998-12-07
EARLIER FILING DATE: 1999-03-24
EARLIER FILING DATE: 1999-03-24
EARLIER FILING DATE: 1998-12-10
EARLIER FILING DATE: 1998-12-10
EARLIER FILING DATE: 1999-03-29
NUMBER OF SEQ ID NOS: 45
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 1
LENGTH: 534
TYPE: PRT
ORGANISM: Aspergillus niger
FEATURE:
NAME/KEY: SIGNAL
LOCATION: (1)...(24)
US-09-455-679-1

Query Match 100.0%; Score 2771; DB 4; Length 534;
Best Local Similarity 100.0%; Pred. No. 1,1e-231;
Matches 534; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSFSLALSGVCTGLANVISKRATLDLWSLNEATVARTALINNIGDGMVSGADSGI 60
Db 1 MSFSLALSGVCTGLANVISKRATLDLWSLNEATVARTALINNIGDGMVSGADSGI 60
Qy 61 VVASPTDNPPYFYTWTRDSGLVLTLYDLFRNGDTSLSLTIENTYISAQAIYQGISNPSG 120
Db 61 VVASPTDNPPYFYTWTRDSGLVLTLYDLFRNGDTSLSLTIENTYISAQAIYQGISNPSG 120
Qy 121 DLSSGAGLGEPEKFNVDETAYTGSWGRPQRDGPALRATAMIGFGQWLLDNGYSTATDIYW 180
Db 121 DLSSGAGLGEPEKFNVDETAYTGSWGRPQRDGPALRATAMIGFGQWLLDNGYSTATDIYW 180
Qy 181 PLVRNDLSYVAQYNNQTYDLMEEVNGSSFPTIAVQHRALVEGSAFATVAGSSCSWCDSQ 240
Db 181 PLVRNDLSYVAQYNNQTYDLMEEVNGSSFPTIAVQHRALVEGSAFATVAGSSCSWCDSQ 240
Qy 241 APELICYSQFWTGSFILLANFDSRSRGDANTLLGSIHTFPDEAACDSTFQPCSPRALA 300
Db 241 APELICYSQFWTGSFILLANFDSRSRGDANTLLGSIHTFPDEAACDSTFQPCSPRALA 300
Qy 301 NHKEVVDSPRSIYTLNDGLSDSEAVAVGRYPEDTYNGNPMFLCTLAABEQLYDALYQMD 360
Db 301 NHKEVVDSPRSIYTLNDGLSDSEAVAVGRYPEDTYNGNPMFLCTLAABEQLYDALYQMD 360

Qy 361 KQGSLEVTVDLDFPKALYSDPATGTYSSTSSSTYSIVDAVKTFPDGFVSIYETHAASNG 420
Db 361 KQGSLEVTVDLDFPKALYSDPATGTYSSTSSSTYSIVDAVKTFPDGFVSIYETHAASNG 420
Qy 421 SMSSEQYDKSDGEQSLASRLDTSWYALLLTANNRRNSVVPASWGETSASSVPGTCAATSAIG 480
Db 421 SMSSEQYDKSDGEQSLASRLDTSWYALLLTANNRRNSVVPASWGETSASSVPGTCAATSAIG 480
Qy 481 TYSSTVTVSWPSIYAVGTGTTTATPTGSGSVTSKTKTATASKSTTTTRSGMSL 534
Db 481 TYSSTVTVSWPSIYAVGTGTTTATPTGSGSVTSKTKTATASKSTTTTRSGMSL 534

RESULT 5
US-09-351-814-2
Sequence 2, Application US/09351814

Patent No. 6352851
GENERAL INFORMATION:
APPLICANT: Nielsen, Bjarne Roenfeldt
APPLICANT: Svendsen, Allan
APPLICANT: Pedersen, Henrik
APPLICANT: Vind, Jesper
APPLICANT: Hendriksen, Hanne Vang
APPLICANT: Frandsen, Torben Peter
TITLE OF INVENTION: Glucoamyase Variants
FILE REFERENCE: 5636.200-US
CURRENT FILING DATE: 1999-07-12
EARLIER FILING DATE: 1999-07-12
EARLIER FILING DATE: 1998-07-15
EARLIER FILING DATE: 1998-07-15
EARLIER FILING DATE: 1998-12-17
EARLIER FILING DATE: 1998-12-17
EARLIER FILING DATE: 1998-07-21
EARLIER FILING DATE: 1998-07-21
EARLIER FILING DATE: 1999-01-12
NUMBER OF SEQ ID NOS: 81
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 2
LENGTH: 534
TYPE: PRT
ORGANISM: Aspergillus niger
FEATURE:
NAME/KEY: SIGNAL
LOCATION: (1)...(24)
US-09-351-814-2

Query Match 100.0%; Score 2771; DB 4; Length 534;
Best Local Similarity 100.0%; Pred. No. 1,1e-231;
Matches 534; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSFSLALSGVCTGLANVISKRATLDLWSLNEATVARTALINNIGDGMVSGADSGI 60
Db 1 MSFSLALSGVCTGLANVISKRATLDLWSLNEATVARTALINNIGDGMVSGADSGI 60
Qy 61 VVASPTDNPPYFYTWTRDSGLVLTLYDLFRNGDTSLSLTIENTYISAQAIYQGISNPSG 120
Db 61 VVASPTDNPPYFYTWTRDSGLVLTLYDLFRNGDTSLSLTIENTYISAQAIYQGISNPSG 120
Qy 121 DLSSGAGLGEPEKFNVDETAYTGSWGRPQRDGPALRATAMIGFGQWLLDNGYSTATDIYW 180
Db 121 DLSSGAGLGEPEKFNVDETAYTGSWGRPQRDGPALRATAMIGFGQWLLDNGYSTATDIYW 180
Qy 181 PLVRNDLSYVAQYNNQTYDLMEEVNGSSFPTIAVQHRALVEGSAFATVAGSSCSWCDSQ 240
Db 181 PLVRNDLSYVAQYNNQTYDLMEEVNGSSFPTIAVQHRALVEGSAFATVAGSSCSWCDSQ 240
Qy 241 APELICYSQFWTGSFILLANFDSRSRGDANTLLGSIHTFPDEAACDSTFQPCSPRALA 300
Db 241 APELICYSQFWTGSFILLANFDSRSRGDANTLLGSIHTFPDEAACDSTFQPCSPRALA 300
Qy 301 NHKEVVDSPRSIYTLNDGLSDSEAVAVGRYPEDTYNGNPMFLCTLAABEQLYDALYQMD 360
Db 301 NHKEVVDSPRSIYTLNDGLSDSEAVAVGRYPEDTYNGNPMFLCTLAABEQLYDALYQMD 360

```
Db 301 NHKEVDSFRSIYTLNDGLSDSEAVAVGRYPEDTYNGNPMFLCTLAAEQYDALYQMD 360
Qy 361 KQSGLEVTDVSLDFPKALYSDAATGTYSSSSSTYSIVDAVKTFADGFSIVETHAASNG 420
Db 361 KQSGLEVTDVSLDFPKALYSDAATGTYSSSSSTYSIVDAVKTFADGFSIVETHAASNG 420
Qy 421 SMSBQYDKSDGEQLSARDLTWSYAALLTANNRRNSVVPASWGETSASVPGTCAATSAIG 480
Db 421 SMSBQYDKSDGEQLSARDLTWSYAALLTANNRRNSVVPASWGETSASVPGTCAATSAIG 480
Qy 481 TYSSVTVTSWPSIYATGTTTATPTGSGSVTSTSKTTATASKTSTTRSGMSL 534
Db 481 TYSSVTVTSWPSIYATGTTTATPTGSGSVTSTSKTTATASKTSTTRSGMSL 534
```

```
RESULT 6
; Sequence 9, Application US/09821616
; Patent No. 6620924
; GENERAL INFORMATION:
; APPLICANT: Nielsen, Bjarne R.
; APPLICANT: Nielsen, Ruby
; APPLICANT: Lembeck, Jan
; TITLE OF INVENTION: Thermostable Glucoamylase
; FILE REFERENCE: 5279.200-US
; CURRENT APPLICATION NUMBER: US/09/821,616
; PRIOR APPLICATION NUMBER: 2001-03-29
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/199,290
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 1557/97
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-12-30
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 0925/98
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-07-10
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/070,746
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-01-08
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/094,344
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-07-28
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/979,673
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/107,657
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-30
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 9
; LENGTH: 534
; TYPE: PRT
; ORGANISM: Aspergillus niger
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)...(24)
; US-09-821-616-9
```

```
Query Match 100.0%; Score 2771; DB 4; Length 534;
Best Local Similarity 100.0%; Pred. No. 1.1e-211;
Matches 534; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MSFRLALSGVCTGLANVISKRATLDSWLSNEATVARTAILNNIGADGAWVGADSGI 60
Db 1 MSFRLALSGVCTGLANVISKRATLDSWLSNEATVARTAILNNIGADGAWVGADSGI 60
Qy 61 VVASPSTDNPDYFYTWTRDSGLVLTIVDLFRNGDTSLSLTIEINYISAQAIVOGINSNSG 120
Db 61 VVASPSTDNPDYFYTWTRDSGLVLTIVDLFRNGDTSLSLTIEINYISAQAIVOGINSNSG 120
Qy 121 DLSSGAGLGEPRKFNVDETATYTGSGRPPQRDGPALRATAMIGFGOMLLDNGYSTATDIYW 180
Db 121 DLSSGAGLGEPRKFNVDETATYTGSGRPPQRDGPALRATAMIGFGOMLLDNGYSTATDIYW 180
Qy 121 DLSSGAGLGEPRKFNVDETATYTGSGRPPQRDGPALRATAMIGFGOMLLDNGYSTATDIYW 180
Db 121 DLSSGAGLGEPRKFNVDETATYTGSGRPPQRDGPALRATAMIGFGOMLLDNGYSTATDIYW 180
Qy 181 PLYVNDLSYVAQYNNQGYDYLMEEVNCGSFFTIYVQHRALVEGSAFATAVSSSCSWCDSQ 240
Db 181 PLYVNDLSYVAQYNNQGYDYLMEEVNCGSFFTIYVQHRALVEGSAFATAVSSSCSWCDSQ 240
Qy 241 APEILCYLQSFMTGSLFLANFDSRSRGKDNANTLLGSIHTFDPEAACDSDTFQPCSPRALA 300
```

```
Db 241 APEILCYLQSFMTGSLFLANFDSRSRGKDNANTLLGSIHTFDPEAACDSDTFQPCSPRALA 300
Qy 301 NHKEVDSFRSIYTLNDGLSDSEAVAVGRYPEDTYNGNPMFLCTLAAEQYDALYQMD 360
Db 301 NHKEVDSFRSIYTLNDGLSDSEAVAVGRYPEDTYNGNPMFLCTLAAEQYDALYQMD 360
Qy 361 KQSGLEVTDVSLDFPKALYSDAATGTYSSSSSTYSIVDAVKTFADGFSIVETHAASNG 420
Db 361 KQSGLEVTDVSLDFPKALYSDAATGTYSSSSSTYSIVDAVKTFADGFSIVETHAASNG 420
Qy 421 SMSBQYDKSDGEQLSARDLTWSYAALLTANNRRNSVVPASWGETSASVPGTCAATSAIG 480
Db 421 SMSBQYDKSDGEQLSARDLTWSYAALLTANNRRNSVVPASWGETSASVPGTCAATSAIG 480
Qy 481 TYSSVTVTSWPSIYATGTTTATPTGSGSVTSTSKTTATASKTSTTRSGMSL 534
Db 481 TYSSVTVTSWPSIYATGTTTATPTGSGSVTSTSKTTATASKTSTTRSGMSL 534
```

```
RESULT 7
; Sequence 13, Application US/09351814
; Patent No. 6352851
; GENERAL INFORMATION:
; APPLICANT: Nielsen, Bjarne Roenfeldt
; APPLICANT: Svendsen, Allan
; APPLICANT: Pedersen, Henrik
; APPLICANT: Vind, Jesper
; APPLICANT: Hendriksen, Hanne Vang
; APPLICANT: Frandsen, Torben Peter
; TITLE OF INVENTION: Glucoamylase Variants
; FILE REFERENCE: 5636.200-US
; CURRENT APPLICATION NUMBER: US/09/351,814
; PRIOR FILING DATE: 1999-07-12
; PRIOR APPLICATION NUMBER: PA 1998 00937
; PRIOR FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: PA 1998 01667
; PRIOR FILING DATE: 1998-12-17
; PRIOR APPLICATION NUMBER: 60/093,528
; PRIOR FILING DATE: 1998-07-21
; PRIOR APPLICATION NUMBER: 60/115,545
; PRIOR FILING DATE: 1999-01-12
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 13
; LENGTH: 640
; TYPE: PRT
; ORGANISM: ASPERGILLUS NIGER
; US-09-351-814-13
```

```
Query Match 99.0%; Score 2742; DB 4; Length 640;
Best Local Similarity 99.6%; Pred. No. 4.9e-229;
Matches 528; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Qy 1 MSFRLALSGVCTGLANVISKRATLDSWLSNEATVARTAILNNIGADGAWVGADSGI 60
Db 1 MSFRLALSGVCTGLANVISKRATLDSWLSNEATVARTAILNNIGADGAWVGADSGI 60
Qy 61 VVASPSTDNPDYFYTWTRDSGLVLTIVDLFRNGDTSLSLTIEINYISAQAIVOGINSNSG 120
Db 61 VVASPSTDNPDYFYTWTRDSGLVLTIVDLFRNGDTSLSLTIEINYISAQAIVOGINSNSG 120
Qy 121 DLSSGAGLGEPRKFNVDETATYTGSGRPPQRDGPALRATAMIGFGOMLLDNGYSTATDIYW 180
Db 121 DLSSGAGLGEPRKFNVDETATYTGSGRPPQRDGPALRATAMIGFGOMLLDNGYSTATDIYW 180
Qy 121 DLSSGAGLGEPRKFNVDETATYTGSGRPPQRDGPALRATAMIGFGOMLLDNGYSTATDIYW 180
Db 121 DLSSGAGLGEPRKFNVDETATYTGSGRPPQRDGPALRATAMIGFGOMLLDNGYSTATDIYW 180
Qy 181 PLYVNDLSYVAQYNNQGYDYLMEEVNCGSFFTIYVQHRALVEGSAFATAVSSSCSWCDSQ 240
Db 181 PLYVNDLSYVAQYNNQGYDYLMEEVNCGSFFTIYVQHRALVEGSAFATAVSSSCSWCDSQ 240
Qy 241 APEILCYLQSFMTGSLFLANFDSRSRGKDNANTLLGSIHTFDPEAACDSDTFQPCSPRALA 300
```


Db 241 APEILCYLOSFWTGSFILANFPDSSRSKGDANTLLGSIHTFDEPAACDDSTFQPCSPRALA 300
QY 301 NHEKVVDSFRSITYTLNDGSDSEAAVAVGRYPEDTYNGNPMFLCTLLAAEQLYDALYQWD 360
Db 301 NHEKVVDSFRSITYTLNDGSDSEAAVAVGRYPEDTYNGNPMFLCTLLAAEQLYDALYQWD 360
QY 361 KQGSLEVTDVSLDFPKALYSDAATGYSSSSSTYSIYDAVTFPADGPFVSIYETHAASNG 420
Db 361 KQGSLEVTDVSLDFPKALYSDAATGYSSSSSTYSIYDAVTFPADGPFVSIYETHAASNG 420
QY 421 SMSSEYDKSDGSQLSARDLTWSYVALLTANNRNSVVPASWGETSASSVPGTCAATSAIG 480
Db 421 SMSSEYDKSDGSQLSARDLTWSYVALLTANNRNSVVPASWGETSASSVPGTCAATSAIG 480
QY 481 TYSSVTYVSWPSIVATGTTTATPTGSGSVTSTSKTTATASKTSTTSS 530
Db 481 TYSSVTYVSWPSIVATGTTTATPTGSGSVTSTSKTTATASKTSTTSS 530

RESULT 8

US-09-236-063-1
Sequence 1, Application US/09236063
Patent No. 6537792

GENERAL INFORMATION:

APPLICANT: Allen, Martin
APPLICANT: Fang, Tsuei-Yun
APPLICANT: Li, Yuxing
APPLICANT: Liu, Hsuan-Liang
APPLICANT: Chen, Hsui-Mei
APPLICANT: Coutinho, Pedro
APPLICANT: Hanzatko, Richard
APPLICANT: Ford, Clark
TITLE OF INVENTION: PROTEIN ENGINEERING OF GLUCOAMYLASE TO
TITLE OF INVENTION: INCREASE PH OPTIMUM, SUBSTRATE SPECIFICITY AND
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESSES:
ADDRESSER: Kohn & Associates
STREET: 30500 No. 6537792Western Hwy.
CITY: Farmington Hills
STATE: Michigan
COUNTRY: US
ZIP: 48334
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/236,063
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Kohn, Kenneth I.
REGISTRATION NUMBER: 30,955
REFERENCE/DOCKET NUMBER: 0812.00001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (248) 539-5050
TELEFAX: (248) 539-5050
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 616 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEICAL: NO
ORIGINAL SOURCE:
ORGANISM: Aspergillus

Query Match 94.8%; Score 2628; DB 4; Length 616;
Best Local Similarity 99.6%; Pred. No. 3,5e-219;

Matches 504; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 25 ATLDSWLSNEATVARTALLNIGADGAVSGADSGIYVASTNDPDTFYTTTRSGVYL 84
Db 1 ATLDSWLSNEATVARTALLNIGADGAVSGADSGIYVASTNDPDTFYTTTRSGVYL 60
QY 85 KTLVDFPENGSTSLSTIENYISAQAIYQGISNPGDSSGAGLGEPPFNVDETAYTGSW 144
Db 61 KTLVDFPENGSTSLSTIENYISAQAIYQGISNPGDSSGAGLGEPPFNVDETAYTGSW 120
QY 145 GRPQDGPALBATAWIGFGWLLDNGYSTATDIWPLVRNDLSYVAQYMNQGYDLWE 204
Db 121 GRPQDGPALBATAWIGFGWLLDNGYSTATDIWPLVRNDLSYVAQYMNQGYDLWE 180
QY 205 VNGSSFFITIAVQHRALVGSAPATVSGSSCSWCSQAPEILCYLOSFWTGSFILANFPSS 264
Db 181 VNGSSFFITIAVQHRALVGSAPATVSGSSCSWCSQAPEILCYLOSFWTGSFILANFPSS 240
QY 265 RSGKANTLLGSIHTFDEPAACDDSTFQPCSPRALANHEKVVDSFRSITYTLNDGSDSEA 324
Db 241 RSGKANTLLGSIHTFDEPAACDDSTFQPCSPRALANHEKVVDSFRSITYTLNDGSDSEA 300
QY 325 VAVGRYPEDTYNGNPMFLCTLLAAEQLYDALYQWDKQGSLEVTDVSLDFPKALYSDAAT 384
Db 301 VAVGRYPEDTYNGNPMFLCTLLAAEQLYDALYQWDKQGSLEVTDVSLDFPKALYSDAAT 360
QY 385 GTYSSSSSTYSIYDAVTFPADGPFVSIYETHAASNGSSEYDKSDGSQLSARDLTWSYA 444
Db 361 GTYSSSSSTYSIYDAVTFPADGPFVSIYETHAASNGSSEYDKSDGSQLSARDLTWSYA 420
QY 445 ALLTANNRNSVVPASWGETSASSVPGTCAATSAIGTSSVTYVSWPSIVATGTTTAT 504
Db 421 ALLTANNRNSVVPASWGETSASSVPGTCAATSAIGTSSVTYVSWPSIVATGTTTAT 480
QY 505 PTGSGSVTSTSKTTATASKTSTTSS 530
Db 481 PTGSGSVTSTSKTTATASKTSTTSS 506

RESULT 9

US-09-199-290-34
Sequence 34, Application US/09199290
Patent No. 6255084

GENERAL INFORMATION:

APPLICANT: Nielsen, Bjarne R.
APPLICANT: Nielsen, Ruby
APPLICANT: Lembeck, Jan
TITLE OF INVENTION: Thermostable Glucoamylase
FILE REFERENCE: 5279,200-US
CURRENT APPLICATION NUMBER: US/09/199,290
CURRENT FILING DATE: 1998-11-24
EARLIER APPLICATION NUMBER: 1557/97
EARLIER FILING DATE: 1997-12-30
EARLIER APPLICATION NUMBER: 0925/98
EARLIER FILING DATE: 1998-07-10
EARLIER APPLICATION NUMBER: 60/070,746
EARLIER FILING DATE: 1998-01-08
EARLIER APPLICATION NUMBER: 60/094,344
EARLIER FILING DATE: 1998-07-28
EARLIER APPLICATION NUMBER: 08/979,673
EARLIER FILING DATE: 1997-11-26
EARLIER APPLICATION NUMBER: 09/107,657
EARLIER FILING DATE: 1998-06-30
NUMBER OF SEQ ID NOS: 34
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 34
LENGTH: 618
TYPE: PRT
ORGANISM: Talaromyces emersonii
US-09-199-290-34

Query Match 61.0%; Score 1689; DB 3; Length 618;
Best Local Similarity 61.2%; Pred. No. 9,1e-138;

Matches 322; Conservative 75; Mismatches 111; Indels 18; Gaps 5;
QY 7 LALSGLVCTGLANV--ISKRA--LDSWLSNEATVARTALINNIGADGAWYSGADSGIVVA 63
Db 9 LCTIGLPPAFAAPAVARATGSLDSFLATETPIALOGVANNIPGNADVAGASAGIVVA 68
QY 64 SPSTDNDEYFTYTRDGLVLTIVDLFRNGDTLSLTENIYSAQAIYOGISNPSGDL 123
Db 69 SPSPSDNRYFSWTWRDALTAFTKYLVDFAFIAENKPLEQTIQOYISAQKVQTIISNPSGDL 128
QY 124 SGAGLPEKPNVDEATYATGSGWRPQDGPALRATAMIGFQGMILDNNGTSTPATDIWPLV 183
Db 129 TG-GLGPKENVETAFTGWRPQDGPALRATAMIGFQGMILDNNGTSTPATDIWPLV 187
QY 184 RNDLSYAOYWNQOTGYDLMEEVNCSFETTAVOHRALVEGSAFATAVGSSCMWDSQAP 243
Db 188 QNDLSYITQWNSSTFPLMEVEGSSFFTTAVQHRALVEGNALATRLNHTCSNCSQAPQ 247
QY 244 ILCTLOSPWTSFLANF--DSSRGKDANTLGSITHFDPBAACDDSTFQPCSPRALAN 302
Db 248 VLCTLOSPWTSFLANF--DSSRGKDANTLGSITHFDPBAACDDSTFQPCSPRALAN 307
QY 303 KEVVDSPRSITYTLNDGLSDSEAVAVGRYPEDTYNGNPMFLCTLAABOLYDALYQMDK 362
Db 308 KVTYDSFRSITYALNSGIAEGSAVAVGRYPEDTYNGNPMFLCTLAABOLYDALYQMDK 367
QY 363 GSELEVDVSLDFPKALYSDAATGYSSSSSTYSIYDAVKTADGFSIYETHAASGSM 422
Db 368 GSISITDVSLPFQDIPSAVGTGYSSTFNDIISAVQYIGDGYISYEKTPSDGSL 427
QY 423 SEQYDKSDGEOLARDLTWSYALALTANNRNSVVPASWETSASSVPGTCAATSAIGTY 482
Db 428 TEQFSRDTGPLSASALTWSYALALTANNRNSVVPASWETSASSVPGTCAATSAIGTY 487
QY 483 SSVTVTSPSIVATGTTTATPTGSGSVTSTSKTATASKTSTT 528
Db 488 STATNTWPS-----SGSGSSTTSSAPCTTPTPSAVT 520
RESULT 10
US-09-821-616-34
; Sequence 34, Application US/09821616
; Patent No. 6620924
; GENERAL INFORMATION:
; APPLICANT: Nielsen, Bjarne R.
; APPLICANT: Nielsen, Ruby
; APPLICANT: Lehmebeck, Jan
; TITLE OF INVENTION: Thermostable Glucoamylase
; FILE REFERENCE: 5279.200-US
; CURRENT APPLICATION NUMBER: US/09/821,616
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/199,290
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-11-24
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 1557/97
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-12-30
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 0925/98
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-07-10
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/070,746
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-01-08
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/094,344
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-07-28
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/979,673
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/107,657
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-30
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 34
; LENGTH: 618
; TYPE: PRT
; ORGANISM: Talaromyces emersonii
US-09-821-616-34

Query Match 61.0%; Score 1689; DB 4; Length 618;
Best Local Similarity 61.2%; Pred. No. 9.1e-138;
Matches 322; Conservative 75; Mismatches 111; Indels 18; Gaps 5;
QY 7 LALSGLVCTGLANV--ISKRA--LDSWLSNEATVARTALINNIGADGAWYSGADSGIVVA 63
Db 9 LCTIGLPPAFAAPAVARATGSLDSFLATETPIALOGVANNIPGNADVAGASAGIVVA 68
QY 64 SPSTDNDEYFTYTRDGLVLTIVDLFRNGDTLSLTENIYSAQAIYOGISNPSGDL 123
Db 69 SPSPSDNRYFSWTWRDALTAFTKYLVDFAFIAENKPLEQTIQOYISAQKVQTIISNPSGDL 128
QY 124 SGAGLPEKPNVDEATYATGSGWRPQDGPALRATAMIGFQGMILDNNGTSTPATDIWPLV 183
Db 129 TG-GLGPKENVETAFTGWRPQDGPALRATAMIGFQGMILDNNGTSTPATDIWPLV 187
QY 184 RNDLSYAOYWNQOTGYDLMEEVNCSFETTAVOHRALVEGSAFATAVGSSCMWDSQAP 243
Db 188 QNDLSYITQWNSSTFPLMEVEGSSFFTTAVQHRALVEGNALATRLNHTCSNCSQAPQ 247
QY 244 ILCTLOSPWTSFLANF--DSSRGKDANTLGSITHFDPBAACDDSTFQPCSPRALAN 302
Db 248 VLCTLOSPWTSFLANF--DSSRGKDANTLGSITHFDPBAACDDSTFQPCSPRALAN 307
QY 303 KEVVDSPRSITYTLNDGLSDSEAVAVGRYPEDTYNGNPMFLCTLAABOLYDALYQMDK 362
Db 308 KVTYDSFRSITYALNSGIAEGSAVAVGRYPEDTYNGNPMFLCTLAABOLYDALYQMDK 367
QY 363 GSELEVDVSLDFPKALYSDAATGYSSSSSTYSIYDAVKTADGFSIYETHAASGSM 422
Db 368 GSISITDVSLPFQDIPSAVGTGYSSTFNDIISAVQYIGDGYISYEKTPSDGSL 427
QY 423 SEQYDKSDGEOLARDLTWSYALALTANNRNSVVPASWETSASSVPGTCAATSAIGTY 482
Db 428 TEQFSRDTGPLSASALTWSYALALTANNRNSVVPASWETSASSVPGTCAATSAIGTY 487
QY 483 SSVTVTSPSIVATGTTTATPTGSGSVTSTSKTATASKTSTT 528
Db 488 STATNTWPS-----SGSGSSTTSSAPCTTPTPSAVT 520
RESULT 11
US-09-199-290-7
; Sequence 7, Application US/09199290
; Patent No. 6255084
; GENERAL INFORMATION:
; APPLICANT: Nielsen, Bjarne R.
; APPLICANT: Nielsen, Ruby
; APPLICANT: Lehmebeck, Jan
; TITLE OF INVENTION: Thermostable Glucoamylase
; FILE REFERENCE: 5279.200-US
; CURRENT APPLICATION NUMBER: US/09/199,290
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-11-24
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 1557/97
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-12-30
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 0925/98
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-07-10
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/070,746
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-01-08
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/094,344
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-07-28
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/979,673
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/107,657
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-30
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 591
; TYPE: PRT
; ORGANISM: Talaromyces emersonii
US-09-199-290-7

[illegible]

Db 66 ASGIIVASPDKTNDPYWTYTRDSALTFKCVVDFTFNSYDASIQAEIONTYVAQAHLQGV 125

Qy 116 SNPSGDSLSSGAGLGEPEKENVDETAATYTGSMGRPORBGPALRATAMIGFQWLLDNGYSTA 175

Db 126 SNPSGDSLSSGAGLGEPEKENVDETAATYTGSMGRPORBGPALRATAMIGFQWLLDNGYSTA 185

Qy 176 TDIVPVLVRNDLSYVAQYWNQGYDLMEEVNCGSSFFTTAVQHRALVSGAPATAVGSSCS 235

Db 186 SSIWVPYIKNDLAVYAAQ--NNTGFDLMEEVGSSFFFTVANQHRALVEGALATSLGTS 243

Qy 236 WCOSQAEIILCYLOSFWTGS--FTLANPDSRSRCKDANTLLGSIHTDPEAACDSTFOP 293

Db 244 ACSVAQAQILCFLOSFWSSPSGYILAN---STAKDANTLLGSIHTDPEAACDSTFOP 299

Qy 294 CSPALANHEKVDSPRSIYTYLNDGLSDSEAVAVGRYPEDTYNGNPMFLCTLLAAEQLY 353

Db 300 CSPDALANHKVYTDAPFSIYTSINGIAEGSAVAGRYPEDSPFGNPMYINTLLAAEQLY 359

Qy 354 DALYQMDKQSLVETDVSLEDFKALYSDAATGYSSSSSTYSIVDAVKTFADGFSVIVE 413

Db 360 DALYVWKQGSITVTSTSLAFFKDFSSITPGYSSSTYTTLYNLAISAVADGYNMIVA 419

Qy 414 THASNSMSRQKXSGEQLSARDLTWSYVALLTANNRNSVVPASMGFTSASVPGTC 473

Db 420 QYAGTNGSLSEQFSKTNGEPLSAYDLTWSYVALTLAARRAGVVPSPGAASANSVPAQC 479

Qy 474 AATSAIGYSSVTVTSPSIATGTTTATPTGSGVSTSKTTATASKTSTT 528

Db 480 SATSVGSYSATVTSFPP-----SQTTPASTSAGS-SPASTTATATACSTT 527

RESULT 14
US-08-596-300A-7
Sequence 7, Application US/08596300A
Patent No. 5834191

GENERAL INFORMATION:

APPLICANT:

TITLE OF INVENTION: Production of Heterologous Peptides

NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS:

ADDRESSEE: Klauber & Jackson

STREET: 411 Hackensack Avenue

CITY: Hackensack

STATE: New Jersey

COUNTRY: USA

ZIP: 07601

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/596,300A

FILING DATE: 13-FEB-1996

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Jackson Esq., David A.

REGISTRATION NUMBER: 26,742

REFERENCE/DOCKET NUMBER: 1321-1-001

TELECOMMUNICATION INFORMATION:

TELEPHONE: 201 487-5800

TELEFAX: 201 343-1684

TELEX: 133521

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 626 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHEICAL: NO

FRAGMENT TYPE:

US-08-596-300A-7

Query Match 56.9%; Score 1577; DB 2; Length 626;
Best Local Similarity 57.6%; Pred. No. 4,8e-128;
Matches 306; Conservative 85; Mismatches 130; Indels 10; Gaps 6;

Qy 2 SPSLALSGLVCTGLANVISKRAITLDSMLSNREATYAKTIIINNIGADGWAGSADGIV 61

Db 13 AFAVAVGLPPLLEHKRSDIRK-SVDSYIQTETPIAQKNLNCNIGSGRAGASGAV 71

Qy 62 VASPSTNDPPEYFTWTRDSGLVLTVDLFRNG-DTSLSTIENIYSAQAVOGISPSG 120

Db 72 VASPSKSSPDYWTWTRDALVTKLIYDEFTNDYNTLQNTIQAIAAQAQKQGVNSPSG 131

Qy 121 DLSGAGLGEPEKENVDETAATYTGSMGRPORBGPALRATAMIGFQWLLDNGYSTA 180

Db 132 SLNSGAGLGEPEKENVDETAATYTGSMGRPORBGPALRATAMIGFQWLLDNGYSTA 191

Qy 181 PLVRNDLSYVAQYWNQGYDLMEEVNCGSSFFTTAVQHRALVSGAPATAVGSSCSMD 240

Db 192 PIVYNDLAVYAAQYWNQGYDLMEEVNCGSSFFTTAVQHRALVSGAPATAVGSSCSMD 251

Qy 241 APEILCYLOSFWTGS--FTLANPDSRSRCKDANTLLGSIHTDPEAACDSTFOPCSPRAL 299

Db 252 APQILCFQGSFWSNGYITISNFWYNSGKDINSVLTISNFPDPAAGCDVVTTPQCSRAL 311

Qy 300 ANHKEVVDSPRSIYTYLNDGLSDSEAVAVGRYPEDTYNGNPMFLCTLLAAEQLYALYOW 359

Db 312 ANHKVVDSPRSIYTYLNDGLSDSEAVAVGRYPEDTYNGNPMFLCTLLAAEQLYALYOW 370

Qy 360 DKQGSLEVTVSLDFEFKALYSDAATGYSSSSSTYSIVDAVKTFADGFSVIVETHAASN 419

Db 371 KQGSITVTSTSLAFFKDFSSITPGYSSSTYTTLYNLAISAVADGYNMIVA 430

Qy 420 GMSSEQYDKSDGEOQLSARDLTWSYVALLTANNRNSVVPASMGFTSASVPGTCATSAI 479

Db 431 GSLEQFDKSGAPLSYTHLTWSYVALTLAARRAGVVPSPGAASANSVPAQCSASTVA 490

Qy 480 GYSSVTVTSPSIATGTTTATPTGSGS-----VTISKTATATASKT 524

Db 491 GSYATATATSPANLTFPASTTVPPTQTCADHVELVTNEKVTYSYGT 541

RESULT 15
US-08-596-300A-14
Sequence 14, Application US/08596300A
Patent No. 5834191

GENERAL INFORMATION:

APPLICANT:

TITLE OF INVENTION: Production of Heterologous Peptides

NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS:

ADDRESSEE: Klauber & Jackson

STREET: 411 Hackensack Avenue

CITY: Hackensack

STATE: New Jersey

COUNTRY: USA

ZIP: 07601

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/596,300A

FILING DATE: 13-FEB-1996

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Jackson Esq., David A.

REGISTRATION NUMBER: 26,742

REFERENCE/DOCKET NUMBER: 1321-1-001

TELECOMMUNICATION INFORMATION:

TELEPHONE: 201 487-5800

TELEFAX: 201 343-1684

US-08-596-300A-7

TELEX: 133521
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 626 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
FRAGMENT TYPE:
US-08-596-300A-14

Query Match 56.9%; Score 1577; DB 2; Length 626;
Best Local Similarity 57.6%; Pred. No. 4,8e-128;
Matches 306; Conservative 85; Mismatches 130; Indels 10; Gaps 6;

QY 2 SFRSLALSGVCTGLANVSRATLDSWLSNEATVARTAIINIGADGAWVGADSGIV 61
DB 13 AFQAVTGLPDPHEKRHSIIIR-SVDSYIQETETPIAQKILCNIGASGCRASGAGV 71
QY 62 VASPESTNDNPFYTRDROGLVKTLVDFRNG-DTSLSTIENTYISAQAIVOGISNPSG 120
DB 72 VASPEKSSPDYWTYTRDALVTKLIVDEFINDYNTLQNTIQAYAAQAKLQGVSNPSG 131
QY 121 DLSGAGLGEPEKFNDEFAVTSWGRPORDBPALRATAMIGFQWLLDNGYSTATDIW 180
DB 132 SLSNAGLGEPEKFNVDLQOFTGAKGRPQDGPPLRAIALIGYKMLVSNVADTKSIW 191
QY 181 PLVRNDLSVAQYMNQGYDLWEVNGSFFTIAVOHRALVEGSAFATAVGSSCSCWDSQ 240
DB 192 PIVKNDLAVTAQYMNNTGFDLWEVNSSSFTIASHRALVEGSAFAKSVGSSCSACDAI 251
QY 241 APELICVLOSFWTGS-FILANPDSRSRSGKQANTLLGSIHTPEPAACDSTFQCSPRAL 299
DB 252 APOILCFQOOSFWSNSGYIISNFVNYSRSGKDINSVLTSIHFPAGCDVNTFQPCSDRAL 311
QY 300 ANHKEVDSFRSIYTLNDGLSDSEAVAVGRYPEDFYNGNPFELCTIAAEQLYDALYOW 359
DB 312 ANHKVVDSDMR-FMGVNSGRTGAKAAVGRYAEVDYINGNPMYLAITLAAEQLYDAVYW 370
QY 360 DKQGSLEVTVSLDFEKKALYSDAATGYSSSSSTYSIVDAVKTFPADGFVSIVETHAASN 419
DB 371 KKQGSITVTSTLAFKDLVPSVSTGYSSSSSTYTAIINAVTYADGFVDIVAQYTPSD 430
QY 420 GSMSEQYKSDGEQLSARDLTSYALALTANNRNSVVPASWGETSASVPGTCAATSAI 479
DB 431 GSIAEQFDKSDGAPLSATHLTWSYASFLSAAARRAGIVPPSWGAASANSLPGSCASTVA 490
QY 480 GTYSVTVTSWPSIVATGTTT-ATPTGSSG-----VTSTSKTATASKT 524
DB 491 GSIAATATISFPANLTPASTIVTPPTQTCADHEVLTFENEKVTTSYGOT 541

Search completed: June 17, 2004, 17:48:36
Job time : 25 secs

HIS PAGE LEFT BLANK

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2004, 17:42:38 ; Search time 46 Seconds
(without alignments)
3662.755 Million cell updates/sec

Title: US-10-038-723-2
Perfect score: 2771
Sequence: 1 MSFSLSLALSGVCTGLANV.....SKTTATASKTSTTRSGMSL 534

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL 25.*

1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phage.*
10: sp_plant.*
11: sp_todent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_virus.*
16: sp_bacteriopl.*
17: sp_archaeopl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2669.5	96.3	639	3	Q870G8 aspergillus
2	2564.5	92.5	639	3	Q12537 aspergillus
3	1689	61.0	618	3	Q9C1V4 aspergillus
4	1518	54.8	493	3	O59846 aspergillus
5	1400	50.5	620	3	Q12623 humicola gr
6	1394	50.3	579	3	Q12596 corticium r
7	1286	46.4	571	3	Q9P4C5 aspergillus
8	784	28.3	609	3	Q8J0P8 mucor circi
9	728	26.3	579	3	Q727X9 rhizopus or
10	674	24.3	515	3	O8TFE5 aspergillus
11	187	6.7	599	17	Q96Z10 sulfobus
12	185.5	6.7	636	17	Q9H188 thermoplasma
13	181	6.5	1588	16	Q8XDG4 escherichia
14	175	6.3	1236	3	Q9C105 schizosacch
15	174.5	6.3	659	17	Q97BM7 thermoplasma
16	174.5	6.3	990	16	Q83J22 rhizoglia fl

17	173.5	6.3	1713	3	O8TGE1 aspergillus
18	172	6.2	622	17	Q97ZD0 sulfobus
19	171	6.2	3360	16	Q88XB6 aspergillus
20	170.5	6.2	1275	5	Q76602 caenothabdi
21	169.5	6.1	190	3	Q07070 saccharomyc
22	168	6.1	4106	16	Q8XQD2 aspergillus
23	164	5.9	2016	5	Q9H1T0 pleistococcus
24	163.5	5.9	457	5	Q86AK1 dictyostelium
25	163	5.9	1333	16	Q8PD38 xanthomonas
26	161	5.8	1063	16	Q9CH86 lactococcus
27	161	5.8	2283	2	Q8VQ99 staphylococ
28	160.5	5.8	916	5	Q7YZ10 monosiga br
29	160	5.8	615	17	Q973T2 sulfobus
30	160	5.8	888	5	Q25336 leishmania
31	159	5.7	2271	16	Q99QY4 staphylococ
32	158	5.7	1283	3	Q9URU4 schizosacch
33	156.5	5.6	2232	5	Q81FX6 caenothabdi
34	156	5.6	1306	2	Q93N36 pantoea ana
35	154.5	5.6	2275	16	Q8NUJ3 staphylococ
36	153	5.5	1131	3	Q74851 schizosacch
37	153	5.5	1310	16	Q8E473 streptococ
38	152.5	5.5	2117	3	Q96U11 neurospora
39	152	5.5	2014	16	Q7U3X4 synechococ
40	150	5.4	1383	3	Q874K9 candida alb
41	149	5.4	967	3	Q08294 saccharomyc
42	149	5.4	1001	3	Q05164 saccharomyc
43	147	5.3	2230	16	Q7U7J7 synechococ
44	146	5.3	612	17	Q97VY0 sulfobus
45	145	5.2	860	16	Q88T89 lactobacilli

ALIGNMENTS

RESULT 1
Q870G8 PRELIMINARY: PRT: 639 AA.

AC-0870G8: 01-VUN-2003 (TRENBLrel. 24, Created)
DT 01-VUN-2003 (TRENBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Glucoamylase.
OS Aspergillus niger.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes.
OC Eukaryota; Trichocommata; Mitosporic Trichocommata; Aspergillus.
OX NCBI_TaxID=5061;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=T21;
RX MEDLINE=95066018; PubMed=7975554;
DT 01-VUN-2003 (TRENBLrel. 24, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DE Glucoamylase.
OS Aspergillus niger.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes.
OC Eukaryota; Trichocommata; Mitosporic Trichocommata; Aspergillus.
OX NCBI_TaxID=5061;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=T21;
RX MEDLINE=95066018; PubMed=7975554;
DT 01-VUN-2003 (TRENBLrel. 24, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DE Glucoamylase.
OS Aspergillus niger.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes.
OC Eukaryota; Trichocommata; Mitosporic Trichocommata; Aspergillus.
OX NCBI_TaxID=5061;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=T21;
RX MEDLINE=95066018; PubMed=7975554;
DT 01-VUN-2003 (TRENBLrel. 24, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DE Glucoamylase.
OS Aspergillus niger.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes.
OC Eukaryota; Trichocommata; Mitosporic Trichocommata; Aspergillus.
OX NCBI_TaxID=5061;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=T21;
RX MEDLINE=95066018; PubMed=7975554;
DT 01-VUN-2003 (TRENBLrel. 24, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DE Glucoamylase.
OS Aspergillus niger.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes.
OC Eukaryota; Trichocommata; Mitosporic Trichocommata; Aspergillus.
OX NCBI_TaxID=5061;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=T21;
RX MEDLINE=95066018; PubMed=7975554;
DT 01-VUN-2003 (TRENBLrel. 24, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DE Glucoamylase.
OS Aspergillus niger.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes.
OC Eukaryota; Trichocommata; Mitosporic Trichocommata; Aspergillus.
OX NCBI_TaxID=5061;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=T21;
RX MEDLINE=95066018; PubMed=7975554;
DT 01-VUN-2003 (TRENBLrel. 24, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DE Glucoamylase.
OS Aspergillus niger.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes.
OC Eukaryota; Trichocommata; Mitosporic Trichocommata; Aspergillus.
OX NCBI_TaxID=5061;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=T21;
RX MEDLINE=95066018; PubMed=7975554;
DT 01-VUN-2003 (TRENBLrel. 24, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DE Glucoamylase.
OS Aspergillus niger.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes.
OC Eukaryota; Trichocommata; Mitosporic Trichocommata; Aspergillus.
OX NCBI_TaxID=5061;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=T21;
RX MEDLINE=95066018; PubMed=7975554;
DT 01-VUN-2003 (TRENBLrel. 24, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DE Glucoamylase.
OS Aspergillus niger.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes.
OC Eukaryota; Trichocommata; Mitosporic Trichocommata; Aspergillus.
OX NCBI_TaxID=5061;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=T21;
RX MEDLINE=95066018; PubMed=7975554;
DT 01-VUN-2003 (TRENBLrel. 24, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DE Glucoamylase.
OS Aspergillus niger.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes.
OC Eukaryota; Trichocommata; Mitosporic Trichocommata; Aspergillus.
OX NCBI_TaxID=5061;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=T21;
RX MEDLINE=95066018; PubMed=7975554;
DT 01-VUN-2003 (TRENBLrel. 24, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DE Glucoamylase.
OS Aspergillus niger.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes.
OC Eukaryota; Trichocommata; Mitosporic Trichocommata; Aspergillus.
OX NCBI_TaxID=5061;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=T21;
RX MEDLINE=95066018; PubMed=7975554;
DT 01-VUN-2003 (TRENBLrel. 24, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DE Glucoamylase.
OS Aspergillus niger.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes.
OC Eukaryota; Trichocommata; Mitosporic Trichocommata; Aspergillus.
OX NCBI_TaxID=5061;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=T21;
RX MEDLINE=95066018; PubMed=7975554;
DT 01-VUN-2003 (TRENBLrel. 24, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DE Glucoamylase.
OS Aspergillus niger.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes.
OC Eukaryota; Trichocommata; Mitosporic Trichocommata; Aspergillus.
OX NCBI_TaxID=5061;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=T21;
RX MEDLINE=95066018; PubMed=7975554;
DT 01-VUN-2003 (TRENBLrel. 24, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DE Glucoamylase.
OS Aspergillus niger.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes.
OC Eukaryota; Trichocommata; Mitosporic Trichocommata; Aspergillus.
OX NCBI_TaxID=5061;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=T21;
RX MEDLINE=95066018; PubMed=7975554;
DT 01-VUN-2003 (TRENBLrel. 24, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DE Glucoamylase.
OS Aspergillus niger.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes.
OC Eukaryota; Trichocommata; Mitosporic Trichocommata; Aspergillus.
OX NCBI_TaxID=5061;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=T21;
RX MEDLINE=95066018; PubMed=7975554;
DT 01-VUN-2003 (TRENBLrel. 24, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DE Glucoamylase.
OS Aspergillus niger.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes.
OC Eukaryota; Trichocommata; Mitosporic Trichocommata; Aspergillus.
OX NCBI_TaxID=5061;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=T21;
RX MEDLINE=95066018; PubMed=7975554;
DT 01-VUN-2003 (TRENBLrel. 24, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DE Glucoamylase.
OS Aspergillus niger.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes.
OC Eukaryota; Trichocommata; Mitosporic Trichocommata; Aspergillus.
OX NCBI_TaxID=5061;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=T21;
RX MEDLINE=95066018; PubMed=7975554;
DT 01-VUN-2003 (TRENBLrel. 24, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DE Glucoamylase.
OS Aspergillus niger.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes.
OC Eukaryota; Trichocommata; Mitosporic Trichocommata; Aspergillus.
OX NCBI_TaxID=5061;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=T21;
RX MEDLINE=95066018; PubMed=7975554;
DT 01-VUN-2003 (TRENBLrel. 24, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DE Glucoamylase.
OS Aspergillus niger.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes.
OC Eukaryota; Trichocommata; Mitosporic Trichocommata; Aspergillus.
OX NCBI_TaxID=5061;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=T21;
RX MEDLINE=95066018; PubMed=7975554;
DT 01-VUN-2003 (TRENBLrel. 24, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DE Glucoamylase.
OS Aspergillus niger.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes.
OC Eukaryota; Trichocommata; Mitosporic Trichocommata; Aspergillus.
OX NCBI_TaxID=5061;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=T21;
RX MEDLINE=95066018; PubMed=7975554;
DT 01-VUN-2003 (TRENBLrel. 24, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DE Glucoamylase.
OS Aspergillus niger.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes.
OC Eukaryota; Trichocommata; Mitosporic Trichocommata; Aspergillus.
OX NCBI_TaxID=5061;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=T21;
RX MEDLINE=95066018; PubMed=7975554;
DT 01-VUN-2003 (TRENBLrel. 24, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DE Glucoamylase.
OS Aspergillus niger.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes.
OC Eukaryota; Trichocommata; Mitosporic Trichocommata; Aspergillus.
OX NCBI_TaxID=5061;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=T21;
RX MEDLINE=95066018; PubMed=7975554;
DT 01-VUN-2003 (TRENBLrel. 24, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DE Glucoamylase.
OS Aspergillus niger.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes.
OC Eukaryota; Trichocommata; Mitosporic Trichocommata; Aspergillus.
OX NCBI_TaxID=5061;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=T21;
RX MEDLINE=95066018; PubMed=7975554;
DT 01-VUN-2003 (TRENBLrel. 24, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DE Glucoamylase.
OS Aspergillus niger.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes.
OC Eukaryota; Trichocommata; Mitosporic Trichocommata; Aspergillus.
OX NCBI_TaxID=5061;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=T21;
RX MEDLINE=95066018; PubMed=7975554;
DT 01-VUN-2003 (TRENBLrel. 24, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DE Glucoamylase.
OS Aspergillus niger.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes.
OC Eukaryota; Trichocommata; Mitosporic Trichocommata; Aspergillus.
OX NCBI_TaxID=5061;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=T21;
RX MEDLINE=95066018; PubMed=7975554;
DT 01-VUN-2003 (TRENBLrel. 24, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DE Glucoamylase.
OS Aspergillus niger.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes.
OC Eukaryota; Trichocommata; Mitosporic Trichocommata; Aspergillus.
OX NCBI_TaxID=5061;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=T21;
RX MEDLINE=95066018; PubMed=7975554;
DT 01-VUN-2003 (TRENBLrel. 24, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DE Glucoamylase.
OS Aspergillus niger.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes.
OC Eukaryota; Trichocommata; Mitosporic Trichocommata; Aspergillus.
OX NCBI_TaxID=5061;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=T21;
RX MEDLINE=95066018; PubMed=7975554;
DT 01-VUN-2003 (TRENBLrel. 24, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DE Glucoamylase.
OS Aspergillus niger.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes.
OC Eukaryota; Trichocommata; Mitosporic Trichocommata; Aspergillus.
OX NCBI_TaxID=5061;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=T21;
RX MEDLINE=95066018; PubMed=7975554;
DT 01-VUN-2003 (TRENBLrel. 24, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DE Glucoamylase.
OS Aspergillus niger.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes.
OC Eukaryota; Trichocommata; Mitosporic Trichocommata; Aspergillus.
OX NCBI_TaxID=5061;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=T21;
RX MEDLINE=95066018; PubMed=7975554;
DT 01-VUN-2003 (TRENBLrel. 24, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DE Glucoamylase.
OS Aspergillus niger.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes.
OC Eukaryota; Trichocommata; Mitosporic Trichocommata; Aspergillus.
OX NCBI_TaxID=5061;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=T21;
RX MEDLINE=95066018; PubMed=7975554;
DT 01-VUN-2003 (TRENBLrel. 24, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DE Glucoamylase.
OS Aspergillus niger.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes.
OC Eukaryota; Trichocommata; Mitosporic Trichocommata; Aspergillus.
OX NCBI_TaxID=5061;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=T21;
RX MEDLINE=95066018; PubMed=7975554;
DT 01-VUN-2003 (TRENBLrel. 24, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DE Glucoamylase.
OS Aspergillus niger.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes.
OC Eukaryota; Trichocommata; Mitosporic Trichocommata; Aspergillus.
OX NCBI_TaxID=5061;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=T21;
RX MEDLINE=95066018; PubMed=7975554;
DT 01-VUN-2003 (TRENBLrel. 24, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DE Glucoamylase.
OS Aspergillus niger.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes.
OC Eukaryota; Trichocommata; Mitosporic Trichocommata; Aspergillus.
OX NCBI_TaxID=5061;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=T21;
RX MEDLINE=95066018; PubMed=7975554;
DT 01-VUN-2003 (TRENBLrel. 24, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DE Glucoamylase.
OS Aspergillus niger.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes.
OC Eukaryota; Trichocommata; Mitosporic Trichocommata; Aspergillus.
OX NCBI_TaxID=5061;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=T21;
RX MEDLINE=95066018; PubMed=7975554;
DT 01-VUN-2003 (TRENBLrel. 24, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DE Glucoamylase.
OS Aspergillus niger.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes.
OC Eukaryota; Trichocommata; Mitosporic Trichocommata; Aspergillus.
OX NCBI_TaxID=5061;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=T21;
RX MEDLINE=95066018; PubMed=7975554;
DT 01-VUN-2003 (TRENBLrel. 24, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DE Glucoamylase.
OS Aspergillus niger.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes.
OC Eukaryota; Trichocommata; Mitosporic Trichocommata; Aspergillus.
OX NCBI_TaxID=5061;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=T21;
RX MEDLINE=95066018; PubMed=7975554;
DT 01-VUN-2003 (TRENBLrel. 24, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DE Glucoamylase.
OS Aspergillus niger.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes.
OC Eukaryota; Trichocommata; Mitosporic Trichocommata; Aspergillus.
OX NCBI_TaxID=5061;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=T21;
RX MEDLINE=95066018; PubMed=7975554;
DT 01-VUN-2003 (TRENBLrel. 24, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DE Glucoamylase.
OS Aspergillus niger.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes.
OC Eukaryota; Trichocommata; Mitosporic Trichocommata; Aspergillus.
OX NCBI_TaxID=5061;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=T21;
RX MEDLINE=95066018; PubMed=7975554;
DT 01-VUN-2003 (TRENBLrel. 24, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DE Glucoamylase.
OS Aspergillus niger.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes.
OC Eukaryota; Trichocommata; Mitosporic Trichocommata; Aspergillus.
OX NCBI_TaxID=5061;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=T21;
RX MEDLINE=95066018; PubMed=7975554;
DT 01-VUN-2003 (TRENBLrel. 24, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DE Glucoamylase.
OS Aspergillus niger.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes.
OC Eukaryota; Trichocommata; Mitosporic Trichocommata; Aspergillus.
OX NCBI_TaxID=5061;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=T21;
RX MEDLINE=95066018; PubMed=7975554;
DT 01-VUN-2003 (TRENBLrel. 24, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DE Glucoamylase.
OS Aspergillus niger.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes.
OC Eukaryota; Trichocommata; Mitosporic Trichocommata; Aspergillus.
OX NCBI_TaxID=5061;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=T21;
RX MEDLINE=95066018; PubMed=7975554;
DT 01-VUN-2003 (TRENBLrel. 24, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DE Glucoamylase.
OS Aspergillus niger.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes.
OC Eukaryota; Trichocommata; Mitosporic Trichocommata; Aspergillus.
OX NCBI_TaxID=5061;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=T21;
RX MEDLINE=95066018; PubMed=7975554;
DT 01-VUN-2003 (TRENBLrel. 24, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DE Glucoamylase.
OS Aspergillus niger.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes.
OC Eukaryota; Trichocommata; Mitosporic Trichocommata; Aspergillus.
OX NCBI_TaxID=5061;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=T21;
RX MEDLINE=95066018; PubMed=7975554;
DT 01-VUN-2003 (TRENBLrel. 24, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DE Glucoamylase.
OS Aspergillus niger.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes.
OC Eukaryota; Trichocommata; Mitosporic Trichocommata; Aspergillus.
OX NCBI_TaxID=5061;
RN [

DR InterPro: IPR000165; Glyco_hydro_15.
DR InterPro: IPR008928; Glyco_trans_6hp.
DR Pfam: PF00686; CBM_20; 1.
DR Pfam: PF00723; Glyco_hydro_15; 1.
DR PRINTS: PR00736; GLHYDRASE15.
DR ProDom: PD001568; CBD 4; 1.
DR PIRSF: PIRSF01031; GLI-a-glycd SMD; 1.
SO SEQUENCE 639 AA; 68308 MW; 025ADB56B6542B69 CRC64

Query Match	96.3%;	Score 2669.5;	DB 3;	Length 639;
Best Local Similarity	98.1%;	Pred. No. 3.9e-162;		
Matches 521; Conservative	1;	Mismatches 6;	Indels 3;	Gaps 3;

Qy	61	VVASPRTNDNDYYTWTTPRSGVLKTVLFPNGGTSLLSTENTISMAQVQGISNPSG	120
Db	61	VVASPRTNDNDYYTWTTPRSGVLKTVLFPNGGTSLLSTENTISMAQVQGISNPSG	120
Qy	121	DLSSGAGLGEPRKENVDETAVTGSGWCRPQRDGPALRATAMI	180
Db	121	DLSSGAGLGEPRKENVDETAVTGSGWCRPQRDGPALRATAMI	180
Qy	181	PLVRNDLSVYAOVYNTGYDLMEEVNGSSFPIANVOHALVEGSAFATAVSSSCSWCDSQ	240
Db	181	PLVRNDLSVYAOVYNTGYDLM-EVNGSSFFPIAQQHALVEGSAFATAVSSSCSWCDSQ	239
Qy	241	APRILCYLGSFWMGSEFLANFDSRSRSGKANT-LIGSIHTFDPPEAACDSTFQPCSRAL	299
Db	240	APRILCYLGSFWMGSEFLANFDSRSRGAQANTLLIGSIHTFDPPEAACDSTFQPCSRAL	299
Qy	300	ANRKEVYDVSRRSITYTINDGLSDSEAVAGRYPEDTYNGNPMFLCTTLAAEQLDALYQW	359
Db	300	ANRKEVYDVSRRSITYTINDGLSDSEAVAVGRYPEDTYNGNPMFLCTTLAAEQLDALYQW	359
Qy	360	DKGGSLEVTDSLDFPKALYSDAACTGSSSSSTSYIVDAVKTFADGPFVSIVETHAASN	419
Db	360	DKGGSLEVTDSLDFPKALYSD-ATGTYSSSSSTSYIVDAVKTFADGPFVSIVETHAASN	418
Qy	420	GSMSEQYDKSDGBOLSARDULTWSYALALTANNRRNSVVPASWGERTSASVGTCAATSAI	479
Db	419	GSMSEQYDKSDGBOLSARDULTWSYALALTANNRRNVVPSAGWERTSASVGTCAATSAI	478
Qy	480	GTYSSVTVYVSWPISIVATGGTTTTATPTGSGSYTSRSKTTATPAASKTSTTTRS	530
Db	479	GTYSSVTVYVSWPISIVATGGTTTTATPTGSGSYTSRSKTTATPAASKTSTSTSS	529

ID	Q12537	PRELIMINARY;	PRT;	639 AA.
AC	Q12537/			
DT	01-NOV-1996 (TrEMBLrel. 01, Created)			
DT	01-NOV-1996 (TrEMBLrel. 01, last sequence update)			
DT	01-OCT-2003 (TrEMBLrel. 25, last annotation update)			
DE	Glucanamyase precursor (EC 3.2.1.3).			
GN	GLA.			
OS	Aspergillus awamori.			
OC	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;			
OC	Eurotiiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus			
OX	NCBI_TaxId=105351;			
RN	[1] _			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=X-100;			
RA	Diagne I., Pechexonov V.T., Bulat S.A., Firsov I.M.;			
RT	"A glucanamyase gene from Aspergillus awamori X-100: structure,			
RT	allocation, and gene phylogeny.";			
RL	Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; U59303; AAB92927.1; -.			
DR	HSSP; P04064; IGA1.			
DR	GO; GO:0005509; F:calcium ion binding; IEA.			

DR	GO: GO:0004339; F:hydrolyase 1,4-alpha-glucosidase activity; IEA.
DR	GO: GO:0016798; F:hydrolyase activity, acting on glycosyl bonds; IEA.
DR	GO: GO:0005576; F:polysaccharide metabolism; IEA.
DR	Interpro: IPR002044; CBD_4.
DR	Interpro: IPR002048; EF-hand.
DR	Interpro: IPR008291; Glu-a-glycosd_SBD.
DR	Interpro: IPR000165; Glyco_hydro_15.
DR	Interpro: IPR008928; Glyco_trans_6hp.
DR	Pfam: PF006686; PSM_20; 1.
DR	Pfam: PF00723; Glyco_hydro_15; 1.
DR	PRINTS; PR00736; GHYDRASES.
DR	ProDom; PD001568; CBD_4; 1.
DR	PROSITE; PS00018; EF_HAND; 1.
DR	PROSITE; PS00820; GLUTCOMYLASE; 1.
DR	PIRSP; PIRSF001031; Glu-a-glycosd_SBD; 1.
KW	Glycosidase; Hydrolase; Signal.
FT	SIGNAL
FT	CHAIN
FT	SEQUENCE
50	639 AA; 68277 MW; 6793D0637D144ACB CRC64;

Query Match	92.5%;	Score 2564.5;	DB 3;	Length 639;
Best Local Similarity	92.3%;	Pred. No. 2e-15;		
Matches 494; Conservative	16;	Mismatches 14;	Indels 11;	Gaps 3

Qy	1	MSRPSLLALSGVCTGLANVTSKCATLDSMTLSNEATAPALAIINNGALGAVWGSADSG
Db	1	MSRPSLLALSGVCTGLASVTSKCATLDSMTLSNEATAPALAIINNGALGAVWGSADSG
Qy	61	VVASPESTNDPDI FYTWTRDSGLVLKTVLDFRNGDTSLSTIENYISAQAIVQGISNP9
Db	61	VVASPESTNDPDI FYTWTRDSGLVIRKTVLDFRNGDITDLSTIEHYISSQAIIQGVSNP9
Qy	121	DLSSGGLGEPFENVDERTATYGSNGRQOROPALRATAMGFGOMLLDNGYTSATDLY
Db	121	DLSSG-GLGEPFENVDERTATYGSNGRQOROPALRATAMGFGOMLLDNGYTSATEIYV
Qy	181	PIVRNDLSVYAQYMNQTYDLMEEVNGSSPFETIAVQHRALVEGSAPATAVGSSCSWCDS
Db	180	PIVRNDLSVYAQYMNQTYDLMEEVNGSSPFETIAVQHRALVEGSAPATAVGSSCSWCDS
Qy	241	APRILICYLOS F-----WTSQFLLANPQSSSGDANTLGSHTTPPEACCDSTFPOCP
Db	240	APRILICYLOS FSSRSGMTGSIYLANFPD----KQTNLLGASHHTPPEACCDSTFPOCP
Qy	296	PRALANHKKEVDSFPRSIYTLNDGLSDSEBAVAGRYPEPTYNGNPNPFLCTLLAAEQLYD
Db	295	PRALANHKKEVDSFPRSIYTLNDGLSDSEBAVAGRYPEPTYNGNPNPFLCTLLAAEQLYD
Qy	356	LIVOMDKQGLSEVTDVSLDFPFKALYSDAATGYSSSSSTYSIYDAKPTADPFVSIIVET
Db	355	LIVOMDKQGLSELTVDVSLDFPFKALYSGAATGYSSSSSTYSISVAKPTADPFVSIIVET
Qy	416	AASNGSMSEQYKSDGEOQLSARDLTWGSYAALLTANNRRNSVVPASGGEFSASVEGTCA
Db	415	AASNGSLSEQPKSDGDELSARDLTWGSYAALLTANNRRNSVVPASGGEFSASVPOTCA
Qy	476	TSALGTGSSVYTWSPSIVANGCTTTATPPGSGSVTSKTTATATSKSTITTRS 530
Db	475	TSAGGTGSSVYTWSPSIVANGATTTATATYGSGVTSKSTTTATATSKSTITTRS 529

RESULT 3		
09C1V4		
ID 09C1V4	PRELIMINARY;	PRT; 618 AA.
AC 09C1V4;		
DT 01-JUN-2001	(TEMBLrel. 17, Created)	
DT 01-JUN-2001	(TEMBLrel. 17, last sequence update)	
DT 01-OCT-2003	(TEMBLrel. 25, last annotation update)	
DE Glucosylase precursor (BC 3.2.1.3).		
GN GA.		
OS Talaromyces emersonii.		
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes		
OC Eurotiales; Trichocommata; Talaromyces.		

OX NCBI_TaxId=68825;
 [1]
 RP SEQUENCE FROM N.A.
 RA Nielsen B.R., Leimbeck J., Frandsen T.P.;
 RT "Cloning, heterologous expression, and enzymatic characterization of a
 RT chemoestable glucamylase from *Talaromyces emersonii*,"
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ304803; CAC28076.1; -
 DR HSSP: P04064; 1GAT.
 DR GO: GO:0004339; F:glucan 1,4-alpha-glucosidase activity; IEA.
 DR GO: GO:0016798; F:hydrolyase activity, acting on glycosyl bonds; IEA.
 DR GO: GO:0005976; P:polysaccharide metabolism; IEA.
 DR InterPro: IPR002044; CBD 4.
 DR InterPro: IPR008291; Gluc-a-glucad SPD.
 DR InterPro: IPR00165; Glyco_hydro_15.
 DR InterPro: IPR008928; Glyco_trans_6np.
 DR Pfam: PF00686; CBM_20; 1.
 DR Pfam: PF00723; Glyco_hydro_15; 1.
 DR PRINTS: PR00736; GLHYDRLASE15.
 DR ProDom: PD001568; CBD 4; 1.
 DR PIRSF: PIRSF00103; Gluc-a-glucad_SBD; 1.
 DR GlycoStase; Hydrolase; Signal.
 FT SIGNAL 1 20 POTENTIAL.
 FT CHAIN 28 618 GLUCOMYLASE.
 SQ SEQUENCE 618 AA; 65429 MW; C48A034A2C06E5B2 CRC64;
 Query Match 61.0%; Score 1689; DB 3; Length 618;
 Best Local Similarity 61.2%; Pred. No. 1,4e-99;
 Matches 322; Conservative 75; Mismatches 111; Indels 18; Gaps 5;

OY 7 LALSGVCTGLANV-ISKRA--LDSWLSNEATVARTALINNIGADGAMVSGDSIVYA 63
 DB 9 LCILGTPAPAPAPAPAPAPATGSLDSFLATEPPIALQGVANNIGPAGAVAGASGIVYA 68
 OY 64 SESTNDPDYFYTWTRDSGLVLTVDLPFNCGDTSLSTIENYISAQAVQGISNPGDLS 123
 DB 69 SPSRSDPNFYGMTDAALTAKYLVDAFIAQNKMLEQITQYISAAKQVITISNPGDLS 128
 OY 124 SGAGIGEPFNVDENAYTSGMGRPQDGPALRATAMIGGOWLNDNGYSTATDIWPLYR 183
 DB 129 TG-GIGEPFNVENETAFTGPMGRPQDGPALRATAMIAVANYLIDNGEASTDEIWPV 187
 OY 184 RNDLSTVAQYVNTQGYDLMEEVNGSSFFITIAVOHRLVSGSAFATVAGSSGMSQAP 243
 DB 188 QNDLSTIYQYVNSSTFDLMEEVGSSFFITIAVOHRLVSGNALRNLHTCSNCSQAP 247
 OY 244 ILCYLOSFMTGSFIILANF-DSSRSGKDANTLLGSIHTPDPAACDSTFQPCSPRALANH 302
 DB 248 VLCFLQSYWTSGYVLANFGSGRSGKDVNSILGSIHTPDPAACDSTFQPCSPRALANH 307
 OY 303 KEVDSFRSIYTLNDGLDSEAVANGRYPEDTYNGNPMWFLCTLAABOLYALYQMDKQ 362
 DB 308 KVVTSFRSIYALINGIAEGSAVANGRYPEDTYNGNPMWFLCTLAABOLYALYQMDKQ 367
 OY 363 GSELETVDSVLDLDFKALYSDAATGTYSSSSSTYSIVDAVKTADGVSIVETHAASNGSM 422
 DB 368 GSISTIDVSLPFDIYPSAIVGTNSGTTNDIISAQVITGIDGLTVEKTPSDGL 427
 OY 423 SEQYKSDGEOULSARDLTWSYAAALLTANNRRNSVVPASGSETSSASVPCTCATSAIGTY 482
 DB 428 TEQFRTDGTPLASALTLTWSYASLLTASARQSVVPASGESSASVPACATATGY 487
 OY 483 SSVYTVTSMPSIVATGGTTTATPTGSGVSTSKTATASKTSTT 528
 DB 488 STATNTVMPSS-----SGSGSSTTSSAPCTTPTSAVT 520

RESULT 4
 ID 059846 PRELIMINARY; PRT; 493 AA.
 AC 059846;
 DT 01-AUG-1998 (TREMBlrel. 07, Created)
 DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)

DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Glucamylase.
 OS Aspergillus oryzae.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eucotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
 OC NCBI_TaxId=5062;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O-1013;
 RX MEDLINE=9817274; PubMed=9511753;
 RA Hata Y., Ichida H., Ichikawa E., Kawato A., Suginami K., Imayasu S.;
 RT "Nucleotide sequence of an alternative glucamylase-encoding gene
 RT (glab) expressed in solid-state culture of *Aspergillus oryzae*,"
 RL Gene 207:127-134(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O-1013;
 RC STRAIN=O-1013;
 RA Hata Y., Ichida H., Kojima Y., Ichikawa E., Kawato A., Suginami K.,
 RA Imayasu S.;
 RT "Comparison of two glucamylases produced by *Aspergillus oryzae* in
 RT solid-state (koji) and in submerged culture,"
 RL J. ferment. Bioeng. 84:532-537(1997).
 DR EMBL: AB007825; BAA25205.1; -
 DR PIR: PC6503; JC6538.
 DR HSSP: P04064; 1GAT.
 DR GO: GO:0004339; F:glucan 1,4-alpha-glucosidase activity; IEA.
 DR GO: GO:0005976; P:polysaccharide metabolism; IEA.
 DR InterPro: IPR00165; Glyco_hydro_15.
 DR InterPro: IPR008928; Glyco_trans_6np.
 DR Pfam: PF00723; Glyco_hydro_15; 1.
 DR PRINTS: PR00736; GLHYDRLASE15.
 DR PROSITE: PS00820; GLUCOMYLASE; 1.
 SQ SEQUENCE 493 AA; 52401 MW; 3707B3A6F0C66256 CRC64;
 Query Match 54.8%; Score 1518; DB 3; Length 493;
 Best Local Similarity 57.5%; Pred. No. 8.6e-89;
 Matches 280; Conservative 82; Mismatches 122; Indels 2; Gaps 2;

OY 5 SLALSGVCTGLANV-ISKRA--LDSWLSNEATVARTALINNIGADGAMVSGDSIVYA 64
 DB 8 SLNLAAGVAVHSPFIHRRQSDLNFAIQPIALQGVANNIGADGKVEGAAGAIIVAS 67
 OY 65 PSTNDPDYFYTWTRDSGLVLTVDLPFNCGDTSLSTIENYISAQAVQGISNPGDLS 124
 DB 68 PSKSNPDYFYTWTRDSGLVLTVDLPFNCGDTSLSTIENYISAQAVQGISNPGDLS 127
 OY 125 GAGIGEPFNVDENAYTSGMGRPQDGPALRATAMIGGOWLNDNGYSTATDIWPLYR 184
 DB 128 GSGLAEPFNVENETAFTGPMGRPQDGPALRATAMIAVANYLIDNGEASTDEIWPV 187
 OY 185 NDLSTVAQYVNTQGYDLMEEVNGSSFFITIAVOHRLVSGSAFATVAGSSGMSQAP 244
 DB 188 NDLSTVAQYVNTQGYDLMEEVNGSSFFITIAVOHRLVSGSAFATVAGSSGMSQAP 246
 OY 245 LCVLOSFMTGSFIILANFSS-RSGKDANTLLGSIHTPDPAACDSTFQPCSPRALANH 303
 DB 247 LCHLQDFPNNGSAVSNLPTNGSGSLDNTLSIGSIHTPDPAACDSTFQPCSPRALANH 306
 OY 304 EYVDSFRSIYTLNDGLDSEAVANGRYPEDTYNGNPMWFLCTLAABOLYALYQMDKQ 363
 DB 307 LVVDSFRSIYTLNDGLDSEAVANGRYPEDTYNGNPMWFLCTLAABOLYALYQMDKQ 366
 OY 364 SLEVTVDLDFKALYSDAATGTYSSSSSTYSIVDAVKTADGVSIVETHAASNGSM 423
 DB 367 QNVVTEISLTPFKDLSNVTTGSAKSSSAVESLSAFTYADGFIIVQETTPPGGALA 426
 OY 424 BOYKSDGEOULSARDLTWSYAAALLTANNRRNSVVPASGSETSSASVPCTCATSAIGTY 483
 DB 427 EYSDQGTTPVASDLTWSYAAALLTANNRRNSVVPASGSETSSASVPCTCATSAIGTY 486
 OY 484 SVTVTSM 490
 DB 487 TPTVGSW 493

RESULT 5

012623 PRELIMINARY; PRT; 620 AA.
 AC 012623;
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
 DT 01-OCT-2003 (TREMblrel. 25, Last annotation update)
 DE Glucoamylase (EC 3.2.1.3).
 GN GLAI.
 OS Humicola grisea var. thermoides.
 OC Eukaryota; Fungi; Ascomycota; mitosporic Ascomycota; Humicola.
 OX NCBI_TaxID=5528;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Betka R.M., Rey M.W., Thompson S.A., Gray G.L., Carmona C.L.,
 RA Power S.D.;
 RT "Molecular cloning, analysis and expression of the glai gene encoding
 RT a thermostable, raw starch-digesting glucoamylase from the fungus
 RT Humicola grisea var. thermoides";
 RL Submitted (JUN-1993) to the EMBL/GenBank/DBJ databases.
 DR EMBL; M89475; AAA3386.1; -.
 DR HSSP; P04064; IGAI.
 DR GO; GO:0004339; F:glucan 1,4-alpha-glucosidase activity; IEA.
 DR GO; GO:0016798; F:hydrolase activity, acting on glycosyl bonds; IEA.
 DR GO; GO:0005976; P:polysaccharide metabolism; IEA.
 DR InterPro; IPR002044; CBD_4.
 DR InterPro; IPR008291; Glu-a-glycosd_SBD.
 DR InterPro; IPR00165; Glyco_hydro_15.
 DR InterPro; IPR008928; Glyco_trans_6hp.
 DR Pfam; PF00686; CBM_20; 1.
 DR Pfam; PF00723; Glyco_hydro_15; 1.
 DR PRINTS; PR00736; GLHYDLASE15.
 DR ProDom; PD001568; CBD_4; 1.
 DR PROSITE; PS00820; GLUCOMYLASE; 1.
 DR PIRSF; PIRSF01031; Glu-a-glycosd_SBD; 1.
 DR GlycoSite; Hydrolase.
 SQ SEQUENCE 620 AA; 66525 MW; 54214FF67E20BE0A CRC64;

Query Match 50.5%; Score 1400; DB 3; Length 620;
 Best Local Similarity 49.5%; Pred. No. 4e-81;
 Matches 269; Conservative 96; Mismatches 146; Indels 32; Gaps 7;
 QY 2 SFRSLALSGLVCTGL-----ANVTSKATLDSWLSNEATVARTAILNINIGADGAWVSGA 56
 DB 3 TFSKLVLVGSVAVGSLGRPHSSRLQERAAVDPIINTEKPLAMNKLANNIPNGKAAPGA 62
 QY 57 DSGIVVAPSTDNPDYFYTWRDGLVLTVD-LFRNGDTSLSTIENYISAQAIYOGI 115
 DB 63 AAGVIVASPSRTDPPYFTWPPDALVLTGIIESLGHNYNTTL-----QQV 108
 QY 116 SNBEGDLSGAGLGEPEKRVNDETAYTSGWRPORDGAPLRATAMIGFQWLLDNGYSTA 175
 DB 109 SNBPGTRADGSLGEBAKFNVDLTAFETGWRPORDGAPLRATAMIGFQWLLDNGYSTA 168
 QY 176 TDIWPLVLRNDLSVYAOYMNQGYDLMEEVNSSFPTIAVOHRLVSGAFAATVAGSSCS 235
 DB 169 KSVWMPVVKNDLATAQYMNETGFDLMEEVNSSFPTIAVSHRALTBGAYIAAQLDTECP 228
 QY 236 WCDQAPEILCYLOSFWT--GSFILANDSS--RSGKDANTLLGSINTFDEBAACDSTF 291
 DB 229 PCTVAPQVLCFQQAFAFNKSNVYVSTGAGEYRSGKANSILASIHNFDEACDNLTF 288
 QY 292 QPCSPRALANKEVVDVFRSLYTLNDGLSDEAVAVGRIPEDTYVNGPMLCTLAAEQ 351
 DB 289 QPCSPRALANKEVVDVFRSLYTLNDGLSDEAVAVGRIPEDTYVNGPMLCTLAAEQ 348
 QY 352 LYDLAYOMDKGSLVETDVSLLDFKALYSDAATGYSSSSSTYSSIVYAVTPADGFPSI 411
 DB 349 LYDLAYOMDKGSLVETDVSLLDFKALYSDAATGYSSSSSTYSSIVYAVTPADGFPSI 408
 QY 412 VETHAASNGSMEQYDKSDGEQLSARDLTSVYALLTANNRRNSVVPASW-GETSASSVP 470

RESULT 6

012596 PRELIMINARY; PRT; 579 AA.
 AC 012596;
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
 DT 01-OCT-2003 (TREMblrel. 25, Last annotation update)
 DE Glucoamylase G2 (EC 3.2.1.3).
 OS Corticium rolfsii.
 OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
 OC Aphyllophorales; Corticiaceae; Corticium.
 OX NCBI_TaxID=39291;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Nagasaka Y., Muraki N., Kimura A., Suto M., Yokota A., Tomita F.;
 RT "Cloning of Corticium rolfsii glucoamylase cDNA and its expression in
 RT Saccharomyces cerevisiae";
 RL Appl. Microbiol. Biotechnol. 44:451-458(1995).
 DR EMBL; D49448; BAA08436.1; -.
 DR HSSP; P04064; IGAI.
 DR GO; GO:0004339; F:glucan 1,4-alpha-glucosidase activity; IEA.
 DR GO; GO:0016798; F:hydrolase activity, acting on glycosyl bonds; IEA.
 DR GO; GO:0005976; P:polysaccharide metabolism; IEA.
 DR InterPro; IPR002044; CBD_4.
 DR InterPro; IPR008291; Glu-a-glycosd_SBD.
 DR InterPro; IPR00165; Glyco_hydro_15.
 DR InterPro; IPR008928; Glyco_trans_6hp.
 DR Pfam; PF00686; CBM_20; 1.
 DR Pfam; PF00723; Glyco_hydro_15; 1.
 DR PRINTS; PR00736; GLHYDLASE15.
 DR ProDom; PD001568; CBD_4; 1.
 DR PIRSF; PIRSF01031; Glu-a-glycosd_SBD; 1.
 DR GlycoSite; Hydrolase.
 SQ SEQUENCE 579 AA; 61548 MW; FF56D83F9D1A9BC9 CRC64;

Query Match 50.3%; Score 1394; DB 3; Length 579;
 Best Local Similarity 51.2%; Pred. No. 8.8e-81;
 Matches 294; Conservative 77; Mismatches 151; Indels 52; Gaps 11;
 QY 3 FRSLLALSGLVCTGLANVTSKATLDSWLSNEATVARTAILNINIGADGAWVSGDGIYV 62
 DB 2 FRSLLALA---ACAVASVSAQSASATAYLTKEASVAKNGVLCNIGSCSGSEGAYSIGVI 58
 QY 63 ASPESTONPDYFYTWRDGLVLTVDLFRNG-DTSLSTIENYISAQAIYOGISNPSGD 121
 DB 59 ASPEKTSIPDVLVYTWTRKSSLVFKMLIDQYTNGLDTLTKTLIDEFVSEATITQOTSNSG 118
 QY 122 LSSGAGLGEPEKRVNDETAYTSGWRPORDGAPLRATAMIGFQWLLDNGYSTATDIWMP 181
 DB 119 VSTG-GLGEPEKRVNDETAYTSGWRPORDGAPLRATAMIGFQWLLDNGYSTATDIWMP 177
 QY 182 LVNRDLISVYAOYMNQGYDLMEEVNSSFPTIAVOHRLVSGAFAATVAG--SSCSWDS 239
 DB 178 IIRKLDLVNYSMDNQITFDLMEEVNSSFPTIAVOHRLVSGAFAATLIGQTSASASTYSA 237
 QY 240 QAPEILCYLOSFW-TGSFILANPDSSRSGKDANTLLGSINTFDEBAACDSTPOPCSPRA 298
 DB 238 TAPSIILCYLOSFW-TGSFILANPDSSRSGKDANTLLGSINTFDEBAACDSTPOPCSDVA 297

```
QY 299 LANHKEVDSFRSIYTLNDGLSDSEAVAGRRYPEDTYNGNPMFLCTLLAAEQLYDALYQ 358
DB 298 LANLKYVDSFRSIYTLNDGLSDSEAVAGRRYPEDTYNGNPMFLCTLLAAEQLYDALYQ 357
QY 359 WDKQSLVETDVSLDFKALYSDAATGTYSSTSSSTYSIVDAVKTFADGFSIVETHAAS 418
DB 358 WKAABEANTVSLAFQOPDSSITAGTYAASSSVTSLISDIOAFDEFDVIVAKYTS 417
QY 419 NGSMEQVYKSDGEQSLASDRLTWSYALLTANNRRNSVVPASMGESASVGTCTA--AT 476
DB 418 SFLSEBQYDKSTGADQSANLTMSYAAALTAAYQARNFTGASMG--AAGVSTSCSTGAT 474
QY 477 SAIGTYSVTVT-----SWPSIVATG-----GTTTAT-----PTGSGSVST 514
DB 475 SFGGSSGSVEVTFDYATTVYGNITITDVSSELGWMTPANGVALLSANYPTWSATIALP 534
QY 515 SKTT-----ATASKTSTTTRSG 531
DB 535 ADTTIQYKYNIDGSTVIMEDALSNREITTPASG 568

RESULT 7
Q9P4C5 PRELIMINARY; PRT; 571 AA.
AC Q9P4C5;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Glucoamylase.
OS Lentinula edodes (Shiitake mushroom) (Lentinus edodes).
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
OC Agaricales; Tricholomataceae; Lentinula.
OX NCBI_TaxID=5353;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20292826; PubMed=10831434;
RA Zhao J., Chen Y.H., Kwan H.S.;
RT "Molecular Cloning, Characterization, and Differential Expression of a
RT Glucoamylase Gene from the Basidiomycetous Fungus Lentinula edodes.";
RL Appl. Environ. Microbiol. 66:2531-2535 (2000).
DR EMBL; AF20541; AAF75523.1; -.
DR HSSP; P04064; IGAI.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0004339; F:glucan 1,4-alpha-glucosidase activity; IEA.
DR GO; GO:0005976; P:polysaccharide metabolism; IEA.
DR InterPro; IPR002044; CBD_4.
DR InterPro; IPR008291; Glu-a-glycd SBD.
DR InterPro; IPR00165; Glyco_hydro_15.
DR InterPro; IPR008928; Glyco_trans_6hp.
DR Pfam; PF00686; CBM_20; 1.
DR Pfam; PF00723; Glyco_hydro_15; 1.
DR PRINTS; PR00736; GLHYDRASE15.
DR ProDom; PD001568; CBD_4; 1.
DR PIRSF; PIRSF001031; Glu-a-glycd SBD; 1.
SQ SEQUENCE 571 AA; 61161 MW; A1BD1ECC848E97FF CRC64;

Query Match 46.4%; Score 1286; DB 3; Length 571;
Best Local Similarity 47.9%; Pred. No. 6.7e-74;
Matches 266; Conservative 91; Mismatches 150; Indels 48; Gaps 10;
```

```
DB 193 TYDLMEEIYSSFFSTAVQHRLRGCTTLGRALIGQTSILASSGNGADFLCTLGQSYMNT 252
QY 254 GSFTLANPDSRSGKQDANTLGSITFEPPEACDDSTPOGSPRALLANHKEVDSFRSIY 313
DB 253 G-YMTANNGGGRSGIDANSVLASIHFTPDAAAGCDAITFQPCSDVALLNFTYVNAFRNAY 311
QY 314 TLNDGLSDSEAVAGRRYPEDTYNGNPMFLCTLLAAEQLYDALYQMDQKSLVETDVSLD 373
DB 312 EINSIGSTNEAVLTRYPEVDYMGNPWLTLLAAEQLYDLSLIVNSGSLDVTNTISLP 371
QY 374 PFKALYSDAATGTYSSTSSSTYSIVDAVKTFADGFSIVETHAASNGSMSEQYDSDGEO 433
DB 372 PFKQFDSITTAGTYSSTSPSEFTLLAAKYTFSDGFVALVAKYTPSSGILAEQISNGAP 431
QY 434 LSARDLWSYALLTANNRRNSVVPASMG-----ETSSASVPG- 471
DB 432 VSAVDLTWSYASGALFAPARAGVIPGSGAQGLSLNCGGPVAQAVSVTFNDASTLEGO 491
QY 472 ---TCAATSAIGTY-----SSVTYVSWPSIVATGTTT-----ATPTGSGSVST 515
DB 492 NYLITGAVDALDEWSTDNAIILSSANYPVWSVTVLPSTDVQYRYIKKDSGTVTWESD 551
QY 516 ---KTTATASKTSTT 527
DB 552 PMMEITTPANGTYAT 566

RESULT 8
Q8J0P8 PRELIMINARY; PRT; 609 AA.
AC Q8J0P8;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Glucoamylase Glam.
OS Mucor circinellus.
OC Eukaryota; Fungi; Zygomycota; Zygomycetes; Mucorales; Mucoraceae;
OC Mucor.
OX NCBI_TaxID=36080;
RN [1]
RP SEQUENCE FROM N.A.
RX Submitted (OCT-2002) to the EMBL/Genbank/DBJ databases.
RA Hougen-larsen J., Pedersen P.A.;
RT EMBL; AY168303; AAN85206.1; -.
DR GO; GO:0004339; F:glucan 1,4-alpha-glucosidase activity; IEA.
DR GO; GO:0005976; P:polysaccharide metabolism; IEA.
DR InterPro; IPR005036; CBM_21.
DR InterPro; IPR008928; Glyco_hydro_15.
DR InterPro; IPR00165; Glyco_trans_6hp.
DR Pfam; PF03370; CBM_21; 1.
DR Pfam; PF00723; Glyco_hydro_15; 1.
DR PRINTS; PR00736; GLHYDRASE15.
DR ProSITE; PS00820; GLUCOAMYLASE; 1.
SQ SEQUENCE 609 AA; 64778 MW; 56D54CD371CC7C1 CRC64;

Query Match 28.3%; Score 784; DB 3; Length 609;
Best Local Similarity 39.1%; Pred. No. 8e-42;
Matches 175; Conservative 73; Mismatches 158; Indels 42; Gaps 11;
```

```

Db      342 LMEBVGNGVHTLWMVRGVLKCANFATRNMGDSTRATYTTAASIKTKIDSFPNSNGQY 401
QY      257 ILANFD-----SSRGKANTL-----LGIHTFDEAACDDSTPCPCPRALANKEVYDS 308
Db      402 VSVSQSVTGVGSKAGYASVLIASNLGSL-----QDCFFYPSGDKMLATVAALAESEK 452
QY      309 FRSIYTLNDGLSDSEAVAVAGRYPEDTY-----YNGNPWFLLCTLAABQLYDALYQMDKOG 363
Db      453 PASIYISINQNLNGILGALIGRYPEDTYNGNNGSQGNPFICTNAPALYIYALIKEMFNNG 512
QY      364 SLEVTVDLDFPKALYSDAATGT-YSSSSSTYSSIVDAVKTFFADGCVSIYETHAANGSM 422
Db      513 GVTYTSISLNFKKFSSAAAGTKYVTGTSFNSLVQNVAAADAFSTIKFHAATNGSM 572
QY      423 SEQYKSDGEOLSARDLTWGYAALLTAN 450
Db      573 SEQYGRADGLMTGARDLTWHSALISAS 600

```

RESULT 9

```

O727X9
ID      O727X9      PRELIMINARY;      PRT;      579 AA.
AC      O727X9;
DT      01-OCT-2003 (TREMBlrel. 25, Created)
DT      01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DE      Glucoamylase (Fragment).
OS      Rhizopus oryzae (Rhizopus delemar).
OC      Eukaryota; Fungi; Zygomycota; Zygomycetes; Mucorales; Mucoraceae;
OX      Rhizopus.
RN      NCBI_TaxID=64495;
RP      SEQUENCE FROM N.A.
RA      Chen G., Yu X.C., Jiang H.Z., Li M.G.;
RL      Submitted (AUG-2003) to the EMBL/Genbank/DBJ databases.
DR      EMBL; AY359821; AAQ18643.1; -.
FT      NON TER      1      26
FT      CHAIN      27      85      glucoamylase.
FT      CHAIN      86      411      glucoamylase.
FT      CHAIN      411      515
SQ      SEQUENCE 579 AA; 62151 MW; 75811D6936BB8751 CRC64;

Query Match      26.3%; Score 728; DB 3; Length 579;
Best Local Similarity 37.1%; Pred. No. 2.8e-38;
Matches 166; Conservative 70; Mismatches 170; Indels 42; Gaps 10;

QY      25 ATLDSSLNEATVARTAILNNGADGAWVSGDSGIIVASPTDNPDYFTWTRDGLV 84
Db      143 STISSWIKKQEGISRFAMLRINP-----PGSATGFIAASLSTAGPDYVYAMTRDAALTS 197
QY      85 KTLVDLRR--NGDTSLSLTIENTYISAQAIYOGISNPGDSSGAGGEPKFNVDERTAYT 141
Db      198 NVIYEYNTTILSGKKTILNVLKDYTVESVKTOSTYVCN-----CLGEPRFNPDSGCT 251
QY      142 GSWGRPORDGALPATAMIGFGQ-WILDNGYTSTATDIPVLVRNDLSYVQYNNQGYD 200
Db      252 GAKWRPQNDGAREATTFILFADGVLTOTKASVYTGTLKRAIFKDLDYVNVWMSGCLD 311
QY      201 LMEVNGSSPFTIAVQHRALVEGSAFATAVGSS--CSWCDQAPBEILCYLOSFWTGSFIL 258
Db      312 LMEVNGVGHFTYLLWVRKGLLGADFAKRNDSSTASTASTANKISSFWVSSNNW 371
QY      259 ANFPS-----SRGKQDANTL-----LGIHTFDEAACDDSTPCPCPRALANKEVYDS 308
Db      372 IQVSQSVTGVGSKKGLDVSTLLAANLGSV-----DDGFPTPGSKKILATVAVAVDS 422
QY      309 FRSIYTLNDGLSDSEAVAVAGRYPEDTY-----YNGNPWFLLCTLAABQLYDALYQMDKOG 363
Db      423 PASIYIPINKULPSYIGNSIGRYPEDTYNGNNGSQGNPFIFLAGYALYIYALIKEMINGC 482
QY      364 SLEVTVDLDFPKALYSDAATGT-TYSSSSSTYSSIVDAVKTFFADGCVSIYETHAANGSM 422

```

```

Db      483 GVTYTSISLDPFKFSSAATSGKKYVTGTSDFNNALQNLIALADRFLSTVQLHAHNGSL 542
QY      423 SEQYKSDGEOLSARDLTWGYAALLTAN 450
Db      543 ABEFRTTGTGARDLTWHSALISAS 570

```

RESULT 10

```

O8TFE5
ID      O8TFE5      PRELIMINARY;      PRT;      515 AA.
AC      O8TFE5;
DT      01-JUN-2002 (TREMBlrel. 21, Created)
DT      01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT      01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE      Glucoamylase precursor (EC 3.2.1.3).
GN      Glu 0111.
OS      Saccharomycopsis fibuligera (Yeast).
OC      Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OX      Saccharomycetales; Saccharomycopsidaceae; Saccharomycopsis.
RN      NCBI_TaxID=4944;
RP      SEQUENCE FROM N.A.
RC      STRAIN=IFO 0111.
RA      Hostinova E., Solovicova A., Janacek S., Gasperik J.;
RT      "Raw starch degrading glucoamylase from Saccharomycopsis fibuligera:
RL      molecular cloning and expression in yeast."
DR      EMBL; AJ111587; CAC83969.1; -.
DR      GO; GO:0004339; F:glucan 1,4-alpha-glucosidase activity; IEA.
DR      GO; GO:0016798; F:hydrolase activity, acting on glycosyl bonds; IEA.
DR      GO; GO:0005976; P:polysaccharide metabolism; IEA.
DR      InterPro; IPR00165; Glyco_hydro_15.
DR      InterPro; IPR008928; Glyco_trans_6np.
DR      Pfam; PF00723; Glyco_hydro_15; 1.
DR      PRINTS; PR00736; GLHTRLASL5.
DR      PROSITE; PS00820; GLUCOAMYLASE; 1.
KW      Signal; Hydrolase; Glycosidase.
FT      SIGNAL      1      26
FT      CHAIN      27      515      GLUCOAMYLASE.
SQ      SEQUENCE 515 AA; 57423 MW; A2F27D9EEA842AE0 CRC64;

Query Match      24.3%; Score 674; DB 3; Length 515;
Best Local Similarity 33.7%; Pred. No. 6.7e-35;
Matches 163; Conservative 82; Mismatches 177; Indels 62; Gaps 16;

QY      24 RÄTLDSLSNEATVARTAILNNGADGAWVSGDSGIIVASPTDNPDYFTWTRDGLV 83
Db      41 KSNFTQIHQOPAVSWYLLQNLIDYPEGQFSAKRGVAVASPTSEPDYFTQWTRDAIT 100
QY      84 LKTLVDLFRN--GDTSLSTIENTYISAQAIYOGISNPGDLS--GAGLGEPRFNVDET 138
Db      101 FLSLIAVEDEHSFENTTLAKVVEYISNTYTLQVSNPSGNGFDPNHDGGEPRFNVDIT 160
QY      139 AYTSWGRPORDGALPATAMIGF-----GQMLL--DNQYT-STADIVPLVRND 186
Db      161 AYTASWGRPQNDGALPATAYVISRILNAVAKNNKELLLAGONGCPYSSASDIYKIIKPD 220
QY      187 LSYVAQYMNQGYVLMREAVNGSSPFTIAVQHRALVEGSAFATAVGSS--CSWCDQAPBEI 244
Db      221 LQHVSTHMSYSGFPLMEBNGSTHFETALVQKALSYGIPLSKTYNDGQFTSMLEKQDAL 280
QY      245 LCYLOSFWTGSFILAND-----SSRGKQDANTLGLIHTFDEPAAACDDSTPCPCPR 296
Db      281 NSYINS---SGFVNSGKHIVESPOLSRGLDSTATYIALITHD---IGDDTYTFPN- 333
QY      297 PALANKEVYDSF-----RSIYTLNDGLSDSEAVAVAGRYPEDTY-----YNGNPWFLLC 344
Db      334 ---VDNSYVLSLTYLLVDNNKRYKINGNY--KAGAAVAGRYPEBYVNGVTSBGNPQOLA 388
QY      345 TLAABQLYDALYQMDK-QGSLEVTVDLDFPKALYSDA--TGTYS-----SSST 393
Db      389 TAYAGQFFYTLAYNSLQKNKLVLEKLVLDLYNSFIADLSKIDSASKOSLTLTYGSDN 448

```

QY 334 YSSIDAVKTFADGFSIVETHAASNGMSSEQYDKSDGEQLSARDLTWSYALLTANNR 453
 DB 449 YGNVAKSLIQFDSFLKVLHDIDNGQLTEENIRYTGQAGAVSLTWSGSLSANR 508
 QY 454 NSV 457
 DB 509 NGLI 512

RESULT 11

Q96210 PRELIMINARY; PRT; 599 AA.
 AC Q96210;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE Hypothetical protein ST2017.
 GN ST2017.
 OS Sulfolobus tokodaii.
 OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
 OC Sulfolobus.
 CX NCBI_Taxid=11955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=JCM 10545 / 7;
 RX MEDLINE=21456156; PubMed=11572479;
 RA Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
 Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
 Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kuchida N., Oguchi A.,
 Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,
 Oshima T., Kikuchi H.;
 RA "Complete genome sequence of an aerobic thermoacidophilic
 RT Crenarchaeon, Sulfolobus tokodaii strain7";
 RL DNA Res. 8:123-140(2001).
 DR EMBL; AP000988; BAB67116.1; -;
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0004339; F:glucan 1,4-alpha-glucosidase activity; IEA.
 DR GO; GO:0004872; F:receptor activity; IEA.
 DR GO; GO:0005215; F:transporter activity; IEA.
 DR GO; GO:0005976; P:polysaccharide metabolism; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR000165; Glyco_hydro_15.
 DR InterPro; IPR008928; Glyco_trans_6hp.
 DR InterPro; IPR000531; TonB_boxc.
 DR Pfam; PF00723; Glyco_hydro_15; 1.
 DR PROSITE; PS00430; TONB_DEPENDENT_REC_1; 1.
 DR Hypothetical protein; Complete proteome.
 SQ SEQUENCE 599 AA; 69661 MW; F8B94CAE731D119 CRC64;

Query Match 6.7%; Score 187; DB 17; Length 599;
 Best Local Similarity 22.3%; Pred. No. 0.001;
 Matches 102; Conservative 58; Mismatches 151; Indels 146; Gaps 23;
 QY 29 SWL-----SNEATVARTAIL-----NNIGADGAWVSGADSGIVVSPSTNDYFYTW 78
 DB 235 AMLVAVARYDLSVRSLLIIAHMQN-----GALPALDPTDIRRNKXT-----YTVWHR 286
 QY 79 DSGVLKTLVDLFRNGDTSLSTIENYISAQIVOGISNPSGDLSSGAGLGEPRKENVDET 138
 DB 287 DAAF-----ASIALTLGY-----QDPINLNFMTKPLFNGFLPQ-KYTC- 327
 QY 139 AYTGWKRP-----QRDPAALATAMIGGOWMLDNGYSTATDIWPLVRDL 187
 DB 338 ---GWSGTWHPWNRSLPIOD-----ETALMLALWVHFSRFTD---IDFVRPLVAPV 377
 QY 188 SYVAQY-----NOTG-----YDLMEVNGSSFTIAVOHRLVVGSAFATAVGSSGMC 237
 DB 378 KKIAFLVSYNDEEGLPLPSYDLMEERLGHFTFTSLAVYAGLMAKYFAEFGBE-NLK 436
 QY 238 D---SQAPEILCYLOSFWTGS-FILANFDSRSRK--DANTLLGSIHTFPDEACDDSTF 291
 DB 437 DKYLTAAEVKKGLERFYGVGHFARTIYEDNSIDKTVASTLFAAIL-----GPF 486

QY 292 QPSRALANKEVVDSPRSITYLNDGLDSEAVAGRYPEPTYX-----NGNPFLECTLA 347
 DB 487 DPXDPVRSNKRKYBEK-----NING-----IAIYENDWYLKODEKSANAFITLIM 534
 QY 348 AAEQLYDALYOWDKGSLVTDVSLDFPKALYSDAATGYSSSSSTYSYIDAVKTFADG 407
 DB 535 LAQY-----ILEGNKKA 550
 QY 408 FVSIYETPAASNGMSSEQYDKSDGEQLSARDLTWSYA 444
 DB 551 YIDWVISHMLPTGILPEQVSPKN-TYPSVAPLWMSHA 586

RESULT 12

Q9HL88 PRELIMINARY; PRT; 636 AA.
 AC Q9HL88;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE Hypothetical protein Ta0342.
 GN Ta0342.
 OS Thermoplasma acidophilum.
 OC Archaea; Euryarchaeota; Thermoplasmales; Thermoplasmales;
 OC Thermoplasma.
 CX NCBI_Taxid=2303;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DSM 1728;
 RX MEDLINE=20479972; PubMed=11029001;
 RA Ruepp A., Graml W., Santos-Martinez M.-L., Koretke K.K., Volker C.,
 Mewes H.-W., Frishman D., Stocker S., Lupas A.N., Baumeister W.;
 RA "The genome sequence of the thermoacidophilic scavenger Thermoplasma
 RT acidophilum";
 RL Nature 407:508-513(2000).
 DR EMBL; AL445064; CAC11486.1; -;
 DR GO; GO:0004339; F:glucan 1,4-alpha-glucosidase activity; IEA.
 DR GO; GO:0005976; P:polysaccharide metabolism; IEA.
 DR InterPro; IPR000165; Glyco_hydro_15.
 DR InterPro; IPR008928; Glyco_trans_6hp.
 DR Pfam; PF00723; Glyco_hydro_15; 1.
 DR Hypothetical protein; Complete proteome.
 SQ SEQUENCE 636 AA; 73269 MW; FFA4656CDB59AFD CRC64;

Query Match 6.7%; Score 185.5; DB 17; Length 636;
 Best Local Similarity 19.0%; Pred. No. 0.0014;
 Matches 86; Conservative 56; Mismatches 139; Indels 171; Gaps 15;
 QY 57 DSGIVVAPSTND-----NPDYFTWTRDSGLVLTVDLFRNGDTS--LSTIENYISAQ 109
 DB 281 DLGALVASCDDSDILKLSHDGYVWPBRDASMAAYAL-SISGISETARFPALMEDSLSE 339
 QY 110 AIVOGISNPSGDLSSGAGLGEPRKENVDETAYTGSWGRQDPGAL-----BATAWIGFGQ 164
 DB 340 GYLHAKNVDDKIAS-----SWLPHVWNGKSTYPIQDEBTALVVAL 381
 QY 165 WLLDNGYSTATDIWPLVRDLSTVAQYN-----QTGYDLME 204
 DB 382 WEYFRKY-----NDIGFTAYRYRLITRADFMTNFVNDNGLPKPSFDLME 428
 QY 205 VNGSSFTIAVOHRLVVGSAFATAVGSSGMCWSQAPBELICYLOSFWTGSFILANFSS 264
 DB 429 RYGHAYVAVVYALKAASNFANVG-----DPD 458
 QY 265 RSGKDANTLLGSIHTFPEAACDDSTFQPSRALANKE--VVDSPRSITYLNDGLSDS 322
 DB 459 LSEKTEAENAEKNYAFDRFYSEDTGY--ARAIIQKPPFTYDLSALTSVL-FGMDA 514
 QY 323 E-----AAVGRIPEDTYN-----GNPFCTIAAEQLYDA 355
 DB 515 DDPKVISIMORISBDLVWNGVGIGIARVONDRYMEVKDDPSVPGNPIITLIMA-RYYMR 573

QY 356 LYOMDKQSGSLEVTIVSLDFFKALYSDAATGYSSSSSTYSISVDAVKTADGPFVSIYETH 415
DB 574 FGDEFEKAMNL-----IOWKSH 590
QY 416 AASGMSSEQYDKSDGQSLARDTLWSYALL 447
DB 591 RQKSGIFSEQINPYNGEPLSVSPLWSHSEPI 622
RESULT 13
Q8XDG4 PRELIMINARY; PRT; 1588 AA.
AC Q8XDG4;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Putative adhesin (Hypothetical protein).
GN Z5029 OR ECS4480 OR H161.
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. II, Burland V., Mau B., Glaesner J.D.,
Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
Rae J.G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Pocomusis K.,
Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
Weich R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7";
RL Nature 409:529-533 (2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / R1MD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Iehli K., Yokoyama K.,
Han C.-G., Ohtsuda E., Nakayama K., Murata T., Tanaka M., Tobe T.,
Iida T., Takami H., Honda T., Sasaki C., Ogasawara N., Yasunaga T.,
Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22 (2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=R1MD 0509952;
RA Makino K., Yutsudo C.H., Yokoyama K., Kubota Y., Kimura S.,
Shinagawa H.;
RT "O157 specific gene similar to H. influenzae adhesin gene";
RL Submitted (DSC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF005587; AAC58749.1; -
DR EMBL; AF002566; BAB37903.1; -
DR EMBL; AB036416; BAB87814.1; -
DR PIR; A86036; A86036.
DR PIR; H91188; H91188.
DR InterPro; IPR008640; Hep_Hag.
DR InterPro; IPR008635; HTM.
DR InterPro; IPR005594; Yada.
DR Pfam; PF05658; Hep_Hag; 13.
DR Pfam; PF05662; HIM; 12.
DR Pfam; PF03895; Yada; 1.
KM Hypothetical protein; Complete proteome.
SQ SEQUENCE 1588 AA; 160150 MM; B2BA4B06EFP28DEC CRC64;
Query Match 6.5%; Score 181; DB 16; Length 1588;
Best Local Similarity 21.7%; Pred. No. 0.0092;
Matches 136; Conservative 95; Mismatches 217; Indels 178; Gaps 30;
QY 15 TGLANVISKRTLDWSLNEATVARTALNNIGADG-AWVGAGSGIVAVSPSTNDNPF 73
DB 576 TTNANNTSNATNTNTNISN-----LTETVTNLGEBALKM--DKONGVPTAAHGTEETSKI 628

QY 74 YTWTRD-----SGLVLTLYDLFRNGDPTSLSTIENYISAQAIVOGISNPSGDL 122
DB 629 -TNVKGODLTGSGTDAVANGSLKTTNDVAATNTNTNATNTN-----ISNLRETV 677
QY 123 SSGAGLGEF--KENVDETATYGSWGRPQRDPALRATAMFGQMLDNGTSTATIWM 180
DB 678 TN--LGEDALIKMDKXNGVPTAAHG--NNTASKITN-----ILDGVVTATSSPAIN 723
QY 181 PLVRNDL-SYVAQW-----NQTG-----YDLMEEVNCSFFTIAVGRALVEGSAFA 227
DB 724 GSQLYDSSNIAIYFGNANASVNTDGVPTGPTYIGE-----TNTYVGDALAAI--NSFS 777
QY 228 TAVGSSCSWCDSQLPILCYLQSFMTGSFI--LANPDSRSKGA--NTLLGSIHTF--- 280
DB 778 TSLGDALWM-DATAGKRSAGHGTNGDASVITVDAGDISDSSSDAVANGSQLHGVSSVVD 836
QY 281 ---DPEAACDDSTFQCSFRALANKEVYDSFRSI-YTLNDGLSDSEAVANGRIPEPTY 335
DB 837 ALGGGAENVNADGTTAPTYTIANADYDVGDAINALIDTTLDDAL-----LMDADAG 887
QY 336 YNGNPWFICTLAABQLYDALYQMDKQGSLEVTIVSLDFKALYSDAATG----- 385
DB 888 ENG-----AFSAHGKDKTASV--TINVANGAISAASDAINGSLYTTWKYI 933
QY 386 -----TYSSSSSTYSISVDAVKTF-----A 405
DB 934 ADALGGDAENVADGTTAPPTTIANAEYNVGALDALLMDDETANGAGAYNASH 993
QY 406 DGFVSIYETHAANGSNSE-QYDKSDGEQSA-----RDLTWSYAL 446
DB 994 DGRKSIITN--VANGSISESDTDAVNGSQLNATNMIEONTQIINQAGNTDATTYIGENG 1051
QY 447 LTANRNSVVPAPSWGRTSASVPGTCAATSAIGTSSVTVTSVPATGCTTTATPT 506
DB 1052 AGINVTATNDGGLAFNDASAGVGAT-----ALG-YNSVAKGSSVAIGGGSVDVTGI 1105
QY 507 GSGSVTSTKTTTATSKTSTTTRSGM 532
DB 1106 ALGSSSVSSRVIAKGSMDTSTITENG 1131
RESULT 14
Q9C105 PRELIMINARY; PRT; 1236 AA.
AC Q9C105;
ID Q9C105;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Putative glucosylase I (Alpha-1,4-glucan glucosidase), extracellular
starch-degrading enzyme, by similarity to S. cerevisiae STAL, contains
chitinase family signature.
GN SPAB1E7.04C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972h-;
RA Wood V., Rajandream M.A., Barrell B.G., Seeger K., Harris D.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL590605; CAC36921.1; -
DR Genes-Strombe; SPAB1E7.04C; -
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001917; Aminoctans II.
DR InterPro; IPR001223; Glyco hydrc_18.
DR Pfam; PF00704; Glyco hydrc_18_18.
DR PROSITE; PS00599; AA_TRANSFER_CLASS_2; 1.
SQ SEQUENCE 1236 AA; 123387 MM; 5A2D33A30B87CDD8 CRC64;

THIS PAGE LEFT BLANK

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2004, 17:36:33 ; Search time 18 Seconds
(without alignments)
1544.748 Million cell updates/sec

Title: US-10-038-723-2
Perfect score: 2771
Sequence: 1 MSFRLALSGLVCTGLANV.....SKTTATASKTSTTTTSGMSL 534

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2742	99.0	640	1	AMYG ASPNG
2	2620.5	94.6	639	1	AMYG ASPSH
3	2588.5	93.4	639	1	AMYG ASPKA
4	1935	69.8	612	1	AMYG ASPOR
5	1577	55.9	626	1	AMYG NEUCR
6	1321	47.7	616	1	AMYG HORRE
7	781.5	28.2	450	1	AMYG SCHPO
8	725	26.2	604	1	AMYG RHIOF
9	671.5	24.2	519	1	AMYG SACFI
10	664.5	24.0	519	1	AMYG SACPI
11	641.5	22.2	624	1	AMYG ARKAD
12	597	21.5	549	1	AMYG YEAST
13	522	18.8	767	1	AMTH SACDI
14	521	18.8	768	1	AMTH SACDI
15	221.5	8.0	615	1	YGI0 METUA
16	159.5	5.8	1196	1	ICGV_PSEEX
17	159	5.7	1210	1	ICEN_PSEFL
18	158	5.7	670	1	YFEG SCHPO
19	158	5.7	3178	1	YF89 CAEBL
20	157	5.7	1258	1	ICEN ERHME
21	156.5	5.6	540	1	GXU1 ASPAC
22	155.5	5.6	1148	1	ICCK_PSEEX
23	155	5.6	1322	1	ICEN_PSEEX
24	154.5	5.6	1200	1	ICEN_PANAN
25	151.5	5.5	1034	1	ICEN_PSEEX
26	151.5	5.5	1306	1	ICEN_PANAN
27	151.5	5.5	1331	1	MANB CALSA
28	150	5.4	1609	1	FGC2 YEAST
29	147	5.3	507	1	YG46 YEAST
30	142	5.1	1419	1	ALAI CANAL
31	141.5	5.1	644	1	XYND CELFI
32	141	5.1	1802	1	HKRI YEAST
33	140.5	5.1	1567	1	ICEN_XANCT

34	138.5	5.0	600	1	SP96 DICDI
35	138	5.0	797	1	VGIX HSVBR
36	137	4.9	556	1	MSC3 YEAST
37	136	4.9	542	1	CH12 RHIOF
38	136	4.9	995	1	Y109 YEAST
39	135.5	4.9	537	1	P39769 diosophila
40	133.5	4.8	1589	1	PHP DROME
41	132	4.8	827	1	XANP_XAN62
42	130.5	4.7	1537	1	PI01 YEAST
43	130.5	4.7	1746	1	TENA_PIG
44	130	4.7	706	1	PLB2 YEAST
45	129.5	4.7	1367	1	AMTH YEAST

ALIGNMENTS

RESULT 1
AMYG ASPNG STANDARD; PRT; 640 AA.
AC P04064; Q92201; Q99179;
DT 01-NOV-1986 (Rel. 03, Created)
DT 01-NOV-1986 (Rel. 03, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glucoamylase G1 and G2 precursor (EC 3.2.1.3) (Glucon 1,4-alpha-glucosidase) (1,4-alpha-D-glucan glucohydrolase).
GN GLA.
OS Aspergillus niger, and
OS Aspergillus awamori.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_Taxid=5061, 105351;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=A.niger;
RX MEDLINE=84261458; PubMed=6204865;
RA Boel E., Hansen M.T., Hjort I., Hoegh I., Fill N.P.;
RT "Two different types of intervening sequences in the glucoamylase gene from Aspergillus niger.";
RL EMBO J. 3:1581-1585(1984).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=A.niger;
RX MEDLINE=84236105; PubMed=6203744;
RA Boel E., Hjort I., Svensson B., Norris K.E., Fill N.P.;
RT "Glucoamylase G1 and G2 from Aspergillus niger are synthesized from two different but closely related mRNAs.";
RL EMBO J. 3:1097-1102(1984).
RN [3]
RP SEQUENCE OF 25-640, AND COMPARISON OF FORMS G1 AND G2.
RC SPECIES=A.niger;
RX MEDLINE=86136085; PubMed=3081341;
RA Svensson B., Larsen K., Gunnarsson A.;
RT "Characterization of a glucoamylase G2 from Aspergillus niger.";
RL Eur. J. Biochem. 154:497-502(1986).
RN [4]
RP SEQUENCE OF 25-640.
RC SPECIES=A.niger;
RA Svensson B., Larsen K., Svendsen I., Boel E.;
RT "The complete amino acid sequence of the glycoprotein, glucoamylase G1, from Aspergillus niger.";
RL Carlsberg Res. Commun. 48:529-544(1983).
RN [5]
RP SEQUENCE FROM N.A.
RC SPECIES=A.awamori;
RX MEDLINE=85085934; PubMed=6440004;
RA Numborg J.H., Meade J.H., Cole G., Lawyer P.C., McCabe P., Schweikart V., Tal R., Wittman V.P., Flataard J.E., Innis M.A.;
RT "Molecular cloning and characterization of the glucoamylase gene of Aspergillus awamori.";
RL Mol. Cell. Biol. 4:2306-2315(1984).
RN [6]
RP REVISIONS.

RC SPECIES=A.awamori;
 RA Nunberg J.H., Meade J.H., Cole G., Lawler F.C., McCabe P.,
 RA Schweickart V., Tal R., Wittman V.P., Flatgaard J.B., Innis M.A.;
 RL Submitted (FEB-1985) to the EMBL/Genbank/DBJ databases.
 RN [7]
 RP SEQUENCE OF 1-11 FROM N.A.
 RC SPECIES=A.niger; STRAIN=ATCC 10864;
 RX MEDLINE=91168302; PubMed=2076554;
 RA Fowler T., Berka R.M., Ward M.;
 RT "Regulation of the gla gene of *Aspergillus niger*.";
 RL Curr. Genet. 18:537-545(1990).
 RN [8]
 RP CONFORMATION OF O-GLYCOSYLATED REGION.
 RX MEDLINE=92189576; PubMed=1546955;
 RA Williamson G., Belshaw N.J., Williamson M.P.;
 RT "O-glycosylation in *Aspergillus* glucanase. Conformation and role
 in binding."; Biochem. J. 282:423-428(1992).
 RL [9]
 RP ACTIVE SITES, AND MUTAGENESIS.
 RC SPECIES=A.awamori;
 RX MEDLINE=90231978; PubMed=1970434;
 RA Sierks M.R., Ford C., Reilly P.J., Svensson B.;
 RT "Catalytic mechanism of fungal glucanase as defined by mutagenesis
 of Asp176, Glu179 and Glu180 in the enzyme from *Aspergillus*
awamori."; Protein Eng. 3:193-198(1990).
 RL [10]
 RP MUTAGENESIS OF TRP-144.
 RC SPECIES=A.awamori;
 RX MEDLINE=90046622; PubMed=2510150;
 RA Sierks M.R., Ford C., Reilly P.J., Svensson B.;
 RT "Site-directed mutagenesis at the active site Trp120 of *Aspergillus*
awamori glucanase."; Protein Eng. 2:621-625(1989).
 RL [11]
 RP MUTAGENESIS.
 RC SPECIES=A.awamori;
 RX MEDLINE=93165653; PubMed=8433972;
 RA Sierks M.R., Ford C., Reilly P.J., Svensson B.;
 RT "Functional roles and substrate locations of Leu177, Trp178 and Asn182
 of *Aspergillus awamori* glucanase determined by site-directed
 mutagenesis."; Protein Eng. 6:75-79(1993).
 RL [12]
 RP CHARACTERIZATION OF CATALYTIC DOMAIN.
 RX MEDLINE=93277459; PubMed=8503847;
 RA Stoffer B., Frandsen T.P., Busk P.K., Schneider P., Svendsen I.,
 RA Svensson B.;
 RT "Production, purification and characterization of the catalytic
 domain of glucanase from *Aspergillus niger*."; Biochem. J. 292:197-202(1993).
 RL [13]
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 25-495.
 RC SPECIES=A.awamori; STRAIN=Var. X100;
 RX MEDLINE=92406872; PubMed=1527049;
 RA Aleshin A., Glibeov A., Firssov L.M., Honzatko R.B.;
 RT "Crystal structure of glucanase from *Aspergillus awamori* var. X100
 to 2.2-A resolution."; J. Biol. Chem. 267:19291-19298(1992).
 RL [14]
 RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 25-495.
 RC SPECIES=A.awamori; STRAIN=Var. X100;
 RX MEDLINE=94253149; PubMed=8195212;
 RA Aleshin A., Firssov L.M., Honzatko R.B.;
 RT "Refined structure for the complex of acarbose with glucanase from
Aspergillus awamori var. X100 to 2.4-A resolution."; J. Biol. Chem. 269:15631-15639(1994).
 RL [15]
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 25-495.
 RC SPECIES=A.awamori; STRAIN=Var. X100;
 RX MEDLINE=94231577; PubMed=8176747;
 RA Aleshin A., Hoffman C., Firssov L.M., Honzatko R.B.;

RT "Refined crystal structures of glucanase from *Aspergillus awamori*
 var. X100."; J. Mol. Biol. 238:575-591(1994).
 RL [16]
 RP STRUCTURE BY NMR OF 533-640.
 RX MEDLINE=96266494; PubMed=8683599;
 RA Sorimachi K., Jacks A.J., Le Gal-Coeffet M.-F., Williamson G.,
 RA Archer D.B., Williamson M.P.;
 RT "Solution structure of the granular starch binding domain of
 glucanase from *Aspergillus niger* by nuclear magnetic resonance
 spectroscopy."; J. Mol. Biol. 259:970-987(1996).
 RL [17]
 RP STRUCTURE BY NMR OF 533-640.
 RX MEDLINE=97341228; PubMed=9195884;
 RA Sorimachi K., Le Gal-Coeffet M.-F., Williamson G., Archer D.B.,
 RA Williamson M.P.;
 RT "Solution structure of the granular starch binding domain of
Aspergillus niger glucanase bound to beta-cyclodextrin."; Structure 5:647-661(1997).
 RL [18]
 RP CATALYTIC ACTIVITY: Hydrolysis of terminal 1,4-linked alpha-D-
 glucose residues successively from non-reducing ends of the chains
 with release of beta-D-glucose.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=G1;
 CC IsoId=P04064-1; Sequence=Displayed;
 CC Name=G2;
 CC IsoId=P04064-2; Sequence=VSP_000262;
 CC -1- SIMILARITY: Belongs to family 15 of glycosyl hydrolases.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X00548; CAA25219.1; -
 DR EMBL: X00712; CAA25303.1; -
 DR EMBL: X00712; CAA25304.1; -
 DR EMBL: K02465; AAB59296.1; -
 DR EMBL: K02465; AAB59297.1; -
 DR EMBL: X56442; CAA39825.1; -
 DR PIR: A29166; A29166.
 DR PIR: A90986; A90986.
 DR PDB: 1AGW; 30-SEP-94.
 DR PDB: 1GLW; 31-JUL-94.
 DR PDB: 3GLY; 01-NOV-94.
 DR PDB: 1DOG; 30-APR-94.
 DR PDB: 1AC0; 07-JUL-97.
 DR PDB: 1ACZ; 07-JUL-97.
 DR PDB: 1GAH; 17-AUG-96.
 DR PDB: 1GAI; 17-AUG-96.
 DR PDB: 1KUL; 11-JUL-96.
 DR PDB: 1KUM; 11-JUL-96.
 DR GlycoStatedB: P04064; -
 DR InterPro: IPR002044; CBD_4.
 DR InterPro: IPR008291; Glu-a-glcSD_SBD.
 DR InterPro: IPR000165; Glyco_hydro_15.
 DR InterPro: IPR008928; Glyco_trans_6hp.
 DR Pfam: PF00686; CBM_20; 1.
 DR Pfam: PF00723; Glyco_hydro_15; 1.
 DR PRINTS: PRSFP01031; Glu-a-glcSD_SBD; 1.
 DR PRINTS: PRS0736; GLHYDLASE15.
 DR PRODOM: PD001568; CBD_4; 1.
 DR PROSITE: PS00820; GLUCAMYLASE; 1.
 KW Hydrolyase; Glycosidase; Polysaccharide degradation; Glycoprotein;
 KW Alternative splicing; Signal; 3D-structure.
 FT SIGNAL 1 18 POTENTIAL.
 FT PROPEP 19 24
 FT CHAIN 25 640 GLUCAMYLASE G1.

Query Match 99.0%; Score 2742; DB 1; Length 640;
 Best Local Similarity 99.6%; Pred. No. 9.8e-174;
 Matches 528; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSFRLSLALSGVCTGLANVSKRATLDSWLSNEATVARTALINNIGADGAWGADSGI 60
 DB 1 MSFRLSLALSGVCTGLANVSKRATLDSWLSNEATVARTALINNIGADGAWGADSGI 60
 QY 61 VVASPTNDPDPFYWTTRDSGLVLTVDLFRNGDTSLLSTIENYISAQAIYQGISNPSG 120
 DB 61 VVASPTNDPDPFYWTTRDSGLVLTVDLFRNGDTSLLSTIENYISAQAIYQGISNPSG 120
 QY 121 DLSGAGLGEPEKFNDETAFTYSGWGRPQDGPALRATAMIGFGQWLNDGVTSTADIW 180
 DB 121 DLSGAGLGEPEKFNDETAFTYSGWGRPQDGPALRATAMIGFGQWLNDGVTSTADIW 180
 QY 181 PLVRNDLSVVAQYNNQGYDLMEVNGSSFTTIAVQHRALVEGSAFATAVSSCSWCDSQ 240
 DB 181 PLVRNDLSVVAQYNNQGYDLMEVNGSSFTTIAVQHRALVEGSAFATAVSSCSWCDSQ 240
 QY 241 APEILCYLQSFMTGFIILANFDSRSGKANTLGSITFPDEACDDSTFQPCSPRALA 300
 DB 241 APEILCYLQSFMTGFIILANFDSRSGKANTLGSITFPDEACDDSTFQPCSPRALA 300
 QY 301 NHKEVVDSPRSITLTNDGSDSEAVAVGRYPEDTYNGNPMFLCTLAABQLYDALYQMD 360
 DB 301 NHKEVVDSPRSITLTNDGSDSEAVAVGRYPEDTYNGNPMFLCTLAABQLYDALYQMD 360
 QY 361 KQGSLEIVDVSIDPFKALYSDAATGYSSSSSTYSSIVAVTFPADGFSIVETHAASNG 420
 DB 361 KQGSLEIVDVSIDPFKALYSDAATGYSSSSSTYSSIVAVTFPADGFSIVETHAASNG 420
 QY 421 SMSQYDSDGSEQLSARDLTWSYALLTANNRRNSVVPASWETSASVPTCAATSAIG 480
 DB 421 SMSQYDSDGSEQLSARDLTWSYALLTANNRRNSVVPASWETSASVPTCAATSAIG 480
 QY 481 TYSSVTWSPSIYATGTTTATPTGSGSVTSTSKTTATASKTSTTSS 530
 DB 481 TYSSVTWSPSIYATGTTTATPTGSGSVTSTSKTTATASKTSTTSS 530

RESULT 2
 AMYG_ASPIH STANDARD; PRT; 639 AA.
 ID AMYG_ASPIH
 AC P22832;
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 01-JUN-1994 (Rel. 29, Last annotation update)
 DE Glucanase precursor (EC 3.2.1.3) (Glucan 1,4-alpha-glucosidase)
 DE (1,4-alpha-D-glucan glucohydrolase).
 GN GLA.
 OS Aspergillus shirousami.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiales; Trichocommataceae; mitosporic Trichocommataceae; Aspergillus.
 CX NCBI_TaxID=5070;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91182400; PubMed=1368603;
 RA Shibusawa I., Gomi K., Iimura Y., Takahashi K., Tamura G., Hara S.;
 RT "Molecular cloning of the glucanase gene of Aspergillus shirousami
 and its expression in Aspergillus oryzae.";
 RL Agric. Biol. Chem. 54:1905-1914(1990).
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of terminal 1,4-linked alpha-D-
 glucose residues successively from non-reducing ends of the chains
 with release of beta-D-glucose.
 CC -1- SIMILARITY: Belongs to family 15 of glycosyl hydrolases.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

CC entries requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib.ch).
 CC EMBL; D10460; BAA01254.1; -
 DR HSP; P04064; IGA1.
 DR InterPro; IPR002044; CBD 4.
 DR InterPro; IPR008291; Glu-a-glucd_SBD.
 DR InterPro; IPR00165; Glyco_hydro_15.
 DR InterPro; IPR008928; Glyco_trans_6hp.
 DR Pfam; PF00686; CEM_20; 1.
 DR Pfam; PF00723; Glyco_hydro_15; 1.
 DR PRINTS; PIRSF01031; Glu-a-glucd_SBD; 1.
 DR PRODOM; PDOM1568; CBD 4; 1.
 DR PROSITE; PS00820; GLUCOMYLASE; 1.
 KW Hydrolase; Glycosidase; Polysaccharide degradation; Glycoprotein;
 KW Signal.
 FT SIGNAL 1 18 POTENTIAL.
 FT PROPEP 19 24 BY SIMILARITY.
 FT CHAIN 25 639 GLUCOMYLASE.
 FT BINDING 143 143 SUBSTRATE (BY SIMILARITY).
 FT ACT_SITE 199 199 CATALYTIC BASE (BY SIMILARITY).
 FT ACT_SITE 202 202 GENERAL ACID CATALYST (BY SIMILARITY).
 FT ACT_SITE 203 203 INTERACT WITH SUBSTRATES (BY SIMILARITY).
 FT DISULFID 233 236 BY SIMILARITY.
 FT DISULFID 245 472 BY SIMILARITY.
 FT DISULFID 285 293 BY SIMILARITY.
 FT CARBOHYD 194 194 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 205 205 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 418 418 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 639 AA; 68130 MW; E93DAE55ED72326 CRC64;

Query Match 94.6%; Score 2620.5; DB 1; Length 639;
 Best Local Similarity 94.3%; Pred. No. 1e-165;
 Matches 500; Conservative 16; Mismatches 13; Indels 1; Gaps 1;

QY 1 MSFRLSLALSGVCTGLANVSKRATLDSWLSNEATVARTALINNIGADGAWGADSGI 60
 DB 1 MSFRLSLALSGVCTGLANVSKRATLDSWLSNEATVARTALINNIGADGAWGADSGI 60
 QY 61 VVASPTNDPDPFYWTTRDSGLVLTVDLFRNGDTSLLSTIENYISAQAIYQGISNPSG 120
 DB 61 VVASPTNDPDPFYWTTRDSGLVLTVDLFRNGDTSLLSTIENYISAQAIYQGISNPSG 120
 QY 121 DLSGAGLGEPEKFNDETAFTYSGWGRPQDGPALRATAMIGFGQWLNDGVTSTADIW 180
 DB 121 DLSGAGLGEPEKFNDETAFTYSGWGRPQDGPALRATAMIGFGQWLNDGVTSTADIW 180
 QY 181 PLVRNDLSVVAQYNNQGYDLMEVNGSSFTTIAVQHRALVEGSAFATAVSSCSWCDSQ 240
 DB 181 PLVRNDLSVVAQYNNQGYDLMEVNGSSFTTIAVQHRALVEGSAFATAVSSCSWCDSQ 240
 QY 241 APEILCYLQSFMTGFIILANFDSRSGKANTLGSITFPDEACDDSTFQPCSPRALA 300
 DB 241 APEILCYLQSFMTGFIILANFDSRSGKANTLGSITFPDEACDDSTFQPCSPRALA 300
 QY 301 NHKEVVDSPRSITLTNDGSDSEAVAVGRYPEDTYNGNPMFLCTLAABQLYDALYQMD 360
 DB 301 NHKEVVDSPRSITLTNDGSDSEAVAVGRYPEDTYNGNPMFLCTLAABQLYDALYQMD 360
 QY 361 KQGSLEIVDVSIDPFKALYSDAATGYSSSSSTYSSIVAVTFPADGFSIVETHAASNG 420
 DB 361 KQGSLEIVDVSIDPFKALYSDAATGYSSSSSTYSSIVAVTFPADGFSIVETHAASNG 420
 QY 421 SMSQYDSDGSEQLSARDLTWSYALLTANNRRNSVVPASWETSASVPTCAATSAIG 480
 DB 421 SMSQYDSDGSEQLSARDLTWSYALLTANNRRNSVVPASWETSASVPTCAATSAIG 480
 QY 481 TYSSVTWSPSIYATGTTTATPTGSGSVTSTSKTTATASKTSTTSS 530
 DB 481 TYSSVTWSPSIYATGTTTATPTGSGSVTSTSKTTATASKTSTTSS 530

RESULT 3					
AMYG_ASPKA	STANDARD;	PRT;	639 AA.		
ID AMYG_ASPKA	P23176;				
DT 01-NOV-1991	(Rel. 20, Created)				
DT 01-NOV-1991	(Rel. 20, Last sequence update)				
DT 10-OCT-2003	(Rel. 42, Last annotation update)				
DE Glucanase I precursor (EC 3.2.1.3) (Glucan 1,4-alpha-glucosidase)					
DE (1,4-alpha-D-glucan glucohydrolase).					
GN GAI.					
OS Aspergillus kawachi (Aspergillus awamori var. kawachi).					
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;					
OC Eurotiales; Trichocommataceae; mitosporic Trichocommataceae; Aspergillus.					
OX NCBI_TaxID=40384;					
RN (1)					
RP SEQUENCE FROM N.A.					
RA Hayashida S., Nakahara K., Kuroda K., Kuhara S., Fukuda K., Sakaki Y.;					
RT "Structure of the raw-starch-affinity site on the Aspergillus awamori					
RT var. kawachi for localization of the raw-starch-affinity site.";					
RL Agric. Biol. Chem. 53:923-929(1989).					
CC (2)					
CC SEQUENCE OF 494-538.					
RA Hayashida S., Nakahara K., Kuroda K., Miyata T., Iwanaga S.;					
RT "Structure of the raw-starch-affinity site on the Aspergillus awamori					
RT var. kawachi glucanase I molecule."					
RL Agric. Biol. Chem. 53:135-141(1989).					
CC -I- CATALYTIC ACTIVITY: Hydrolysis of terminal 1,4-linked alpha-D-					
CC glucose residues successively from non-reducing ends of the chains					
CC with release of beta-D-glucose.					
CC -I- SIMILARITY: Belongs to family 15 of glycosyl hydrolases.					
CC -----					
CC This SWISS-PROT entry is copyright. It is produced through a collaboration					
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -					
CC the European Bioinformatics Institute. There are no restrictions on its					
CC use by non-profit institutions as long as its content is in no way					
CC modified and this statement is not removed. Usage by and for commercial					
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce					
CC or send an email to license@isb-sdb.ch).					
CC -----					
DR EMBL: J00427; BAA00331.1; -.					
DR PIR: JT0479; JT0479.					
DR GLYCOSUBEDB: P23176; -.					
DR InterPro: IPRO02044; CBD_4.					
DR InterPro: IPRO08291; Glu-a-glcEd_SBD.					
DR InterPro: IPRO00165; Glyco_hydro_15.					
DR InterPro: IPRO08928; Glyco_trans_6hp.					
DR Pfam: PF00683; CEM_20; 1.					
DR Pfam: PF00723; Glyco_hydro_15; 1.					
DR PIRSF: PIRSF001031; Glu-a-glcEd_SBD; 1.					
DR PRINTS: PR00736; GLHYDRLASE15.					
DR ProDom: PD001568; CBD_4; 1.					
DR PROSITE: PS00820; GLUTOMYLASE; 1.					
KW Hydrolase; Glycosidase; Polysaccharide degradation; Glycoprotein;					
KW Signal.					
FT SIGNAL	1	18			
FT PROPEP	19	24			
FT CHAIN	25	639			
FT DOMAIN	494	538	GLUCOMYLASE I.		
FT BINDING	143	143	SUBSTRATE (BY SIMILARITY).		
FT ACT_SITE	199	199	CATALYTIC BASE (BY SIMILARITY).		
FT ACT_SITE	202	202	GENERAL ACID CATALYST (BY SIMILARITY);		
FT ACT_SITE	203	203	INTERACT WITH SUBSTRATES (BY SIMILARITY).		
FT DISULFID	233	236	BY SIMILARITY.		
FT DISULFID	245	472	BY SIMILARITY.		
FT DISULFID	285	293	BY SIMILARITY.		
FT CARBOHYD	194	194	N-LINKED (GLCNAC...) (BY SIMILARITY).		
FT CARBOHYD	418	418	O-LINKED (GLCNAC...) (BY SIMILARITY).		
FT CARBOHYD	464	464	O-LINKED (BY SIMILARITY).		
FT CARBOHYD	466	466	O-LINKED (BY SIMILARITY).		
FT CARBOHYD	467	467	O-LINKED (BY SIMILARITY).		
FT CARBOHYD	475	475	O-LINKED (BY SIMILARITY).		
FT CARBOHYD	476	476	O-LINKED (BY SIMILARITY).		
FT CARBOHYD	482	482	O-LINKED (BY SIMILARITY).		

	FT	CARBOHYD	483	485	O-LINKED	(BY SIMILARITY)	
FT	CARBOHYD	485	485	O-LINKED	(BY SIMILARITY)		
FT	CARBOHYD	487	487	O-LINKED	(BY SIMILARITY)		
FT	CARBOHYD	491	491	O-LINKED	(BY SIMILARITY)		
FT	CARBOHYD	495	495	O-LINKED	(BY SIMILARITY)		
FT	CARBOHYD	498	498	O-LINKED	(BY SIMILARITY)		
FT	CARBOHYD	499	499	O-LINKED	(BY SIMILARITY)		
FT	CARBOHYD	500	500	O-LINKED	(BY SIMILARITY)		
FT	CARBOHYD	501	501	O-LINKED	(BY SIMILARITY)		
FT	CARBOHYD	503	503	O-LINKED	(BY SIMILARITY)		
FT	CARBOHYD	505	505	O-LINKED	(BY SIMILARITY)		
FT	CARBOHYD	507	507	O-LINKED	(BY SIMILARITY)		
FT	CARBOHYD	511	511	O-LINKED	(BY SIMILARITY)		
FT	CARBOHYD	512	512	O-LINKED	(BY SIMILARITY)		
FT	CARBOHYD	513	513	O-LINKED	(BY SIMILARITY)		
FT	CARBOHYD	514	514	O-LINKED	(BY SIMILARITY)		
FT	CARBOHYD	516	516	O-LINKED	(BY SIMILARITY)		
FT	CARBOHYD	517	517	O-LINKED	(BY SIMILARITY)		
FT	CARBOHYD	519	519	O-LINKED	(BY SIMILARITY)		
FT	CARBOHYD	521	521	O-LINKED	(BY SIMILARITY)		
FT	CARBOHYD	523	523	O-LINKED	(BY SIMILARITY)		
FT	CARBOHYD	524	524	O-LINKED	(BY SIMILARITY)		
FT	CARBOHYD	525	525	O-LINKED	(BY SIMILARITY)		
FT	CARBOHYD	526	526	O-LINKED	(BY SIMILARITY)		
FT	CARBOHYD	527	527	O-LINKED	(BY SIMILARITY)		
FT	CARBOHYD	528	528	O-LINKED	(BY SIMILARITY)		
FT	CARBOHYD	529	529	O-LINKED	(BY SIMILARITY)		
FT	CARBOHYD	530	530	O-LINKED	(BY SIMILARITY)		
FT	CARBOHYD	531	531	O-LINKED	(BY SIMILARITY)		
FT	CARBOHYD	533	533	O-LINKED	(BY SIMILARITY)		
FT	CARBOHYD	534	534	O-LINKED	(BY SIMILARITY)		
SQ	SEQUENCE	639 AA;	68271 MW;	E112B31A4DD8D6B	CRC64;		
Query Match 93.4%: Score 2588.5; DB 1; Length 639;							
Best Local Similarity 93.8%: Pred. No. 1,3e-163;							
Matches 497; Conservative 15; Mismatches 17; Indels 1; Gaps 1							
QY	1	MSFRSLALSGVLTGLANVISKRATLDSWLSNEATVAPATLANNIGADGAWGADSGI	60				
DB	1	MSFRSLALSGVLTGLANVISKRATLDSWLSNEATVAPATLANNIGADGAWGADSGI	60				
QY	61	VVASSTNPDPFYFWWRDSDGLVLTVDLPNPGTSLSTENTISAQAIYOGISNPG	120				
DB	61	VVASSTNPDPFYFWWRDSDGLVLTVDLPNPGTSLSTENTISAQAIYOGISNPG	120				
QY	121	DLSSAGLGEPKFNVDETAYTGSWGRPORDEPALATAMIGGQWMLDNGYSTATDIYW	180				
DB	121	DLSSG-GLGEPKFNVDETAYTGSWGRPORDEPALATAMIGGQWMLDNGYSTATDIYW	179				
QY	181	PLVRNDLSVYAQYNNQGYDLMEEVNGSSFPTIAVOHRLVGSAPATAVGSSCSWCDSQ	240				
DB	180	PLVRNDLSVYAQYNNQGYDLMEEVNGSSFPTIAVOHRLVGSAPATAVGSSCSWCDSQ	239				
QY	241	APELICLYOSFMTGSFILANPDSNAGDANTLGSITFPEDEACDDSTFQPCSPRALA	300				
DB	240	APOLICLYOSFMTGSFILANPDSNAGDANTLGSITFPEDEACDDSTFQPCSPRALA	299				
QY	301	NHKEVVDSPRSITYLTDGDSDEAAYGRYPEDTYNGNPMFLCTILAAEQLYDLAYQMD	360				
DB	300	NHKEVVDSPRSITYLTDGDSDEAAYGRYPEDTYNGNPMFQSTILAAEQLYDLAYQMD	359				
QY	361	KQGSLEVDVSLDPEFKALYSDPATGYSSSSGTSYISVDAVCTFADGFSIVETHAASNG	420				
DB	360	KQGSLEITDVLDPFKALYSGAICTYSSSSSTYSISVAVCTFADGFSIVETHAASNG	419				
QY	421	SMSQDYKSDGEQSLARDITMSYALTLANNRNSVPASWGETSASSVPGTCAATSAIG	480				
DB	420	SLSQEPDSDGDELSARDITMSYALTLANNRNSVPPSWGETSASSVPGTCAATSAISG	479				
QY	481	TYSSVTYVSWPSIVATGGTTTATPTGSGSVTSTKTAATASKSTTRS	530				
DB	480	TYSSVTYVSWPSIVATGGTTTATPTGSGSVTSTKTAATASKSTITTS	529				

RESULT 4
 AMYG ASPOR STANDARD; PRT; 612 AA.
 AC P36914;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Glucoamylase precursor (EC 3.2.1.3) (Glucan 1,4-alpha-glucosidase)
 DE (1,4-alpha-D-glucan glucohydrolase).
 GN GLA.
 OS Aspergillus oryzae.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurcitrales; Trichocommataceae; mitosporic Trichocommataceae; Aspergillus.
 OC NCBI_TaxID=5062;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91254744; PubMed=1368680;
 RA Hara Y., Kitamoto K., Gomi K., Kumagai C., Tamura G., Hara S.;
 RT "The glucoamylase cDNA from Aspergillus oryzae: its cloning,
 RT nucleotide sequence, and expression in Saccharomyces cerevisiae";
 RL Agric. Biol. Chem. 55:941-949(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92104497; PubMed=1761224;
 RA Hara Y., Tsuchiya K., Kitamoto K., Gomi K., Kumagai C., Tamura G.,
 RT "Nucleotide sequence and expression of the glucoamylase-encoding gene
 RT (glua) from Aspergillus oryzae.";
 RL Gene 108:145-150(1991).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX STRAIN=RIB 40;
 RA Hara S., Tamura G., Kumagai C., Gomi K., Kitamoto K., Tsuchiya K.,
 RT Hara Y.;
 RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of terminal 1,4-linked alpha-D-
 CC glucose residues successively from non-reducing ends of the chains
 CC with release of beta-D-glucose.
 CC -1- SIMILARITY: Belongs to family 15 of glycosyl hydrolases.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; D01035; BAA00841.1; -;
 DR EMBL; D10698; BAA01540.1; -;
 DR PIR; J01346; J01346.
 DR HSSP; P04064; IGA1.
 DR InterPro; IPR002044; CBD_4.
 DR InterPro; IPR008291; Glu-a-glcad_SBD.
 DR InterPro; IPR000165; Glyco_hydro_15.
 DR InterPro; IPR008928; Glyco_trans_6hp.
 DR Pfam; PF00686; CBM_20; 1.
 DR Pfam; PF00723; Glyco_hydro_15; 1.
 DR PIRSF; PIRSF001031; Glu-a-glcad_SBD; 1.
 DR PRINTS; PR00736; GLHYDRLASE15.
 DR ProDom; PD001568; CBD_4; 1.
 DR PROSITE; PS00820; GLUCOMYLASE_1.
 KM Hydrolyase; Glycosidase; Polysaccharide degradation; Glycoprotein;
 KM Signal.
 FT SIGNAL. 1 19 BY SIMILARITY.
 FT PROPEP 20 25 BY SIMILARITY.
 FT CHAIN 26 612 GLUCOMYLASE.
 FT BINDING 146 146 SUBSTRATE (BY SIMILARITY).
 FT ACT_SITE 202 202 CATALYTIC BASE (BY SIMILARITY).
 FT ACT_SITE 205 205 GENERAL ACID CATALYST (BY SIMILARITY).
 FT ACT_SITE 206 206 INTERACT WITH SUBSTRATES (BY SIMILARITY).
 FT DISULFID 236 239 BY SIMILARITY.

FT DISULFID 248 475 BY SIMILARITY.
 FT DISULFID 288 296 BY SIMILARITY.
 FT CARBOHYD 39 39 N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE 612 AA, 65486 MM, CDTB23B5FA978F97 CRC64;
 Query Match 69.8%; Score 1935; DB 1; Length 612;
 Best Local Similarity 70.3%; Pred. No. 1.9e-120;
 Matches 367; Conservative 59; Mismatches 90; Indels 6; Gaps 4;
 QY 1 MSFRLALISGLVCTGLA--NVISKRATLDSWLSNEATYARPAIILNIGADGAWVSGADS 58
 DB 2 VFSSCCLPALALGSSVLAVQPVLRQATGIDTLWISTEAFNSRQAIILNIGADQSOAGASP 61
 QY 59 GIVVASPTSDNDYVYTMTRDGLVLTLYDFRRGDPSTLSTENYSAOIVGQISNP 118
 DB 62 GVIVASPSKSDPDYVYTMTRDGLVLTLYDFRRGDPSTLSTENYSAOIVGQISNP 121
 QY 119 SGDLSSGAGLGEPKENVDETAYTGSMGRPQRDGPALRATAMIGFQWMLDNGYSTATDI 178
 DB 122 SGALSSG-GLGEPKENVDETAYTGSMGRPQRDGPALRATAMISPEMLVENGHTSIATDL 180
 QY 179 VWPVLRNDLSYVAQYWNQTDLMEEVNGSGFTTAVQHRALVEGSAFATVAGSSCWCDD 238
 DB 181 VWPVVRNDLSYVAQYWSQSGFDLMEEVNGSGFTTAVASHRALVEGSSFAKTVGSSCPYCD 240
 QY 239 SQAPRILCYLGSFMTGSRFLINPDSRSRQKANTLLGSHTEPDAACDDSTFOPCSPRA 298
 DB 241 SQAPVRCYLGSEFMTGSRFLINPDSRSRQKANTLLGSHTEPDAACDDSTFOPCSPRA 300
 QY 299 LANHKEVDSFRSITLNDGLSDSEAVAVAGRPEDTYNGNPFCTLLAAAEQLYDALYQ 358
 DB 301 LANHKEVDSFRSITLNDGLSDSEAVAVAGRPEDTYNGNPFCTLLAAAEQLYDALYQ 360
 QY 359 WDKGSLLEVTVSLDPPFALYSDAATGYSSSSSTYSISVDAVKTTPADGFIIVETHAAS 418
 DB 361 WDKGSLLEVTVSLDPPFALYSDAATGYSSSSSTYSISVDAVKTTPADGFIIVETHAAS 420
 QY 419 NSGSEQYDKSPGEBLSARDLTWYSYALTLNNRNSVYVSWGTSASVGTCAATSA 478
 DB 421 TGSMAEQYTKTDGSGTSARDLTWYSYALTLNNRNSVYVSWGTSASVGTCAATSA 480
 QY 479 IGTYSVTVTSWPSIVANAGT--TTTATPTGSGVTSSTKTT 518
 DB 481 SGTYSVTVTSWPSIVANAGT--TTTATPTGSGVTSSTKTT 521
 RESULT 5
 AMYG NEUCR STANDARD; PRT; 626 AA.
 AC P14804; Q9P5U5; -;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Glucoamylase precursor (EC 3.2.1.3) (Glucan 1,4-alpha-glucosidase)
 DE (1,4-alpha-D-glucan glucohydrolase).
 GN GLA-1 OR B5022.70.
 OS Neurospora crassa.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Sordariomycetiales; Sordariaceae; Neurospora.
 OC NCBI_TaxID=5141;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=74-OR23-1A / FGSC 987;
 RX MEDLINE=94037144; PubMed=8221928;
 RA Stone P.J., Makoff A.J., Parish J.H., Radford A.;
 RT "Cloning and sequence analysis of the glucoamylase gene of Neurospora
 RT crassa.";
 RL Curr. Gene. 24:205-211(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX STRAIN=74-OR23-1A / FGSC 987;
 RX MEDLINE=22542210; PubMed=12655011;
 RA Manhaupt G., Montrone C., Haase D., Mewes H.-W., Algen V.,

RA Hoheisel J.D., Fartmann B., Nyakatura G., Kempen F., Maier J.,
 RA Schulte U.;
 RT "What's in the genome of a filamentous fungus? Analysis of the
 RT Neurospora genome sequence.";
 RT Nucleic Acids Res. 31:1944-1954(2003).
 RN [3]
 RP SEQUENCE OF 36-65.
 RC STRAIN=74-OR23-1A / FGSC 987;
 RA Koh-luar S.I., Parish J.H., Bleasby A.U., Pappin D.J.C., Ainley K.,
 RA Johansen F.E., Radford A.J;
 RT "Exported proteins of *Neurospora crassa*: 1-glucosylase.";
 RT Enzyme Microb. Technol. 11:692-695(1989).
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of terminal 1,4-linked alpha-D-
 CC glucose residues successively from non-reducing ends of the chains
 CC with release of beta-D-glucose.
 CC -1- SIMILARITY: Belongs to family 15 of glycosyl hydrolases.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AL67291; CAA47707.1; -.
 DR EMBL; AL355932; CAB91426.1; -.
 DR HSSP; P04064; 1GAI.
 DR InterPro; IPR002044; CBD_4.
 DR InterPro; IPR008291; Glu-a-glucos_SBD.
 DR InterPro; IPR001651; Glyco_hydro_15.
 DR InterPro; IPR008928; Glyco_trans_6hp.
 DR Pfam; PF00686; CBM_20; 1.
 DR Pfam; PF00723; Glyco_hydro_15; 1.
 DR PIRSF; PIRSF01031; Glu-a-glucos_SBD; 1.
 DR PRINTS; PR00736; GLHYDRLAS15.
 DR PRODOM; PD001568; CBD_4; 1.
 DR PROSITE; PS00820; GLUCOAMYLASE; 1.
 KM Hydrolyase; Glycosidase; Polysaccharide degradation; Glycoprotein;
 KM Signal.
 FT SIGNAL 1 19 POTENTIAL.
 FT PROPEP 20 35 POTENTIAL.
 FT CHAIN 36 626 GLUCOAMYLASE.
 FT BINDING 155 155 SUBSTRATE (BY SIMILARITY).
 FT ACT_SITE 211 211 CATALYTIC BASE (BY SIMILARITY).
 FT ACT_SITE 214 214 GENERAL ACID CATALYST (BY SIMILARITY).
 FT ACT_SITE 215 215 INTERACT WITH SUBSTRATE (BY SIMILARITY).
 FT CARBOHYD 106 106 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 206 206 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 217 217 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 82 82 MISSING (IN REF. 1).
 FT CONFLICT 550 550 A -> R (IN REF. 1).
 FT CONFLICT 560 560 V -> L (IN REF. 1).
 SQ SEQUENCE 626 AA; 66474 MW; 54B5BDBA7A3E349 CMC64;

Query Match 56.9%; Score 1577; DB 1; Length 626;
 Best Local Similarity 57.6%; Pred. 8.6e-97;
 Matches 306; Conservative 85; Mismatches 130; Indels 10; Gaps 6;

QY 2 SFRRLALSLGVCGLANVISKRALDLSMNEATVARTALINNGADGAVSGADSGIV 61
 DB 13 AFQAVLGLPDLHKRHSIDIKR-SVDSYITETPIADKLLCNIGASGRASGAASGV 71
 QY 62 VASSTNDPFFYMTWRDGLVLTGLVDLFRNG-DTSLSTLTENYISAQALVQGISNPSG 120
 DB 72 VASSTKSPDYMTWRDGLVLTGLVDLFRNG-DTSLSTLTENYISAQALVQGISNPSG 131
 QY 121 DLSSGAGLGEKPNVDYATVGSWGRPDGPALRATAMIGFGWLLDNGYSTATDIW 180
 DB 132 SLSSGAGLGEKPNVDYATVGSWGRPDGPALRATAMIGFGWLLDNGYSTATDIW 191
 QY 181 PLVNDLSYVAQYNNQGYDLMEVNGSFFTIATVQRAALVEGAFATVAGSSGSCWCSQ 240

DB 192 PIYKNDLAVTAQYWNNTGFLMEVNSSFTTIAASHRALVEGAFAPAKVSGSSCADAI 251
 QY 241 APEILCYLGSPWGS-FILANFDSRSRGKDANTLGSINTFDPEACDSTFQPCSPRAL 299
 DB 252 APOILCFQGSFWSNGSVIIENFVNRSGKDINSVLTIHNFDPACGDVNTFQPCSDRAL 311
 QY 300 ANHKVVDSPRSIYTLNDGSDSEAVAVGRYPETTYNGNPMWPLCTIAAEQYLDALYQW 359
 DB 312 ANHKVVDSPRSIYTLNDGSDSEAVAVGRYPETTYNGNPMWPLCTIAAEQYLDALYQW 370
 QY 360 DKOSLEVTNVIDLFFALYSDAATGYSSSSSTYSISIVDAVKTFADGFYSIVETHAASN 419
 DB 371 KKQGSITVTSLAFPFDDLPVSSTGYSSSSSTYTHIINAATYAGFVDIYQYTPSD 430
 QY 420 GSNSEQYDKSDGEQLSARDLTWYSYALLTANRRNSVVPASWGETSASVPGTCATSAI 479
 DB 431 GSLAEQFDKDSGAPLSATHTLTWSYASPLSAARBRAGIIPPSWGAASANSIPGSCASTVA 490
 QY 480 GTYSSTVTSWPSIVATGCTTTT-ATPTGSGS-----VTSTSKTTAASKT 524
 DB 491 GSATATATSPFPAWLTTPASTVTPPTGTGCAADHEVLVTENEKVTYSYGT 541

RESULT 6
 AMTG HORRE STANDARD; PRT; 616 AA.
 AC Q03045;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Glucosylase P precursor (EC 3.2.1.3) (Glucan 1,4-alpha-glucosidase)
 DE (1,4-alpha-D-glucan glucosylase).
 GN GAMP.
 OS Hormoconis resinae (Creosote fungus) (Amorphotheca resinae).
 OC Eukaryota; Fungi; Ascomycota; Ascomycota incertae sedis;
 OC Amorphothecaceae; Amorphotheca.
 OX NCBI_TaxID=5101;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 20495;
 RX MEDLINE=93146382; PubMed=1490604;
 RA Joutsen J.V.V., Torkkeli T.K.;
 RT "Glucosylase P gene of *Hormoconis resinae*: molecular cloning,
 RT sequencing and introduction into *Trichoderma reesei*.";
 RN FEMS Microbiol. Lett. 78:237-243(1992).
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 20495;
 RX MEDLINE=93365035; PubMed=8358830;
 RA Vainio A.E.I., Torkkeli H.T., Tuusa T., Aho S.A., Fagerstrom B.R.,
 RA Korhola M.P.;
 RT "Cloning and expression of *Hormoconis resinae* glucosylase P cDNA in
 RT *Saccharomyces cerevisiae*.";
 RL Curr. Genet. 24:38-44(1993).
 RN [3]
 RP SEQUENCE OF 72-76, AND CHARACTERIZATION.
 RA MEDLINE=90338987; PubMed=2116499;
 RA Fagerstrom R., Vainio A.E.I., Suoranta K., Pakula T., Kalkkinen N.,
 RA Torkkeli H.T.;
 RT "Comparison of two glucosylases from *Hormoconis resinae*.";
 RL J. Gen. Microbiol. 136:913-920(1990).
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of terminal 1,4-linked alpha-D-
 CC glucose residues successively from non-reducing ends of the chains
 CC with release of beta-D-glucose.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: Belongs to family 15 of glycosyl hydrolases.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

```

CC or send an email to license@sb-sib.ch).
CC -----
CC EMBL, X68143; CAA48243.1; -.
CC EMBL, X67708; CAA47945.1; -.
CC PIR, S33908; S33908.
CC HSP, P04064; IGAI.
CC InterPro; IPR002044; CBD_4.
CC InterPro; IPR008291; Glu-a-glycd SPD.
CC InterPro; IPR00165; Glyco_hydro_15.
CC InterPro; IPR008928; Glyco_trans_6hp.
CC Pfam; PF0686; CBM_20; 1.
CC Pfam; PF0723; Glyco_hydro_15; 1.
CC PIRSF; PIRSF001031; Glu-a-glycd_SBD; 1.
CC PRINTS; PR00736; GLHYDRLASE15.
CC ProDom; PD001568; CBD_4; 1.
CC PROSITE; PS00820; GLUCOMYGLASE; FALSE_NEG.
CC Hydrolase; Glycosidase; Polysaccharide degradation; Glycoprotein;
CC Signal.
CC CHAIN 1 29
CC SIGNAL 30 616
CC BINDING 149 149
CC ACT_SITE 205 149
CC ACT_SITE 208 208
CC ACT_SITE 209 209
CC CARBOHYD 200 200
CC CARBOHYD 427 427
CC SEQUENCE 616 AA; 66432 MW; B5FDC8EBB152FB CRC64;

Query Match 47.7%; Score 1321; DB 1; Length 616;
Best Local Similarity 48.6%; Pred. No. 6, 8e-80;
Matches 257; Conservative 88; Mismatches 150; Indels 34; Gaps 6;

14 CTGLANVIS-----KRAFLDSWLSNEATVARTALLNNIGADGAVSGADSGIVYA 63
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
8 CAGALSLCSLIAAAPTCLKARDLSFISERAILAQALNNIGPDGSAVPAAGAGFYVA 67
64 SPSTNDPDEFTYWTBDSGLVLTVDLPFNGDPTSLSTIENTYISQAIVQGSINPBGD-L 122
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
68 SSKANPDYFTYWSRDSALTLMKIIDFLGNTTLTITTEQIYHAAVAVQIVNSBSGTL 127
123 SSGAGLGEBEKFNVDETAYTSGWRPORDPALRAFAMIGFGWLLDNGYTSTATDIWPL 182
128 PDGVGLGEBEKFNVDGTRFNGPGRPDGPALRAIALMTYSMWLLKNGQFAEAKTRIPWI 187
133 VANDLSYVAQYNNQGYDLMEEVNGSSPFTTAVQHRALVEGSAFATANGSCSWCDSQAP 242
188 IANDLSYVQYNNQSGFDLMEETVYASSPFTTIONHRAALVEGAQLAHDLVCTGCD-QAP 246
243 EILCYLOSFWTGSFLIANP--DSSRSKQDANTLGSIHFPDEACDDSTPQCPSPRALA 300
247 EYLCLFLOSFWNGKTIYVSNINNVNNGKRTGLGNSILGISTFDIDAYCDSPTLPCHSQSLA 306
301 NKEVVDSEFRSIYTLNDGLSDSEAVAVGRYPEDTYNGNPMFLCTLLAAEQLYDALYQMD 360
307 NFKVLTDPRNLTYTINAGIPEGQVAVGRYADVYMGNPWLITLTAALFYLIDAVAQWK 366
361 KQGSLEVDVSLDFKALYSDAATGYSS--SSSTYSSIVDAVKTFADGFSVIVETHAAS 418
367 ARAHVLIVETSLAFKPDYIPEVTVEYKSGNANSPFAQIMDAVTVAYDSYVAIAEKYIIPS 426
419 NGSMEQYDKSDGEOLASARDLTWSYAALLTANNRNSVVPASMGERSASVSGTCAANSA 478
427 NGSLSERQFRDGTPLSALIDLTWSYAAFLTWSQRRAGQYPSWSGSRNALPPTTCSASST 486
479 IGTYSVTVTWSPIVATGCTTTTATPTGSGSVTSTSKTTATAASKTSTT 527
487 PGIY-----TPATAGAPVNTSSCQVSIFFINATT 517

```

```

DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable glucosylase precursor (EC 3.2.1.3) (Glucan 1,4-alpha-
DE glucosidase) (1,4-alpha-D-glucan glucosylase) (Melotic expression
DE upregulated protein 17).
GN MEU17 OR SPBC14C8.05C.
OS Schizosaccharomyces pombe (Plesion yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=972.
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
RA Gentles S., Godle A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holtroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsle K.,
RA James K., Jones I., Jones M., Leacher S., McDonald S., Mclean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odeil C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinovitch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Vancsteel E., Aert R., Robben J., Grymonpre B.,
RA Wolfgen I., Vansteel E., Rieger M., Schaefer M., Moller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Meisel D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambut R., Purrelle B.,
RA Goffeau A., Cadieu E., Dreno S., Gloux S., Laureau V., Mottier S.,
RA Galbreath F., Aves S.J., Xiang Z., Hunt C., Moore K., Huret S.M.,
RA Lucas M., Rochet M., Gallardin C., Tallada V.A., Garçon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forburg S.L.,
RA Cernutt L., Lowe T., McCombie W.R., Paulsen I., Pochaikin J.,
RA Shpakovski G.V., Useery D., Barrett B.G., Nurse P.;
RA The genome sequence of Schizosaccharomyces pombe.";
RA Nature 415:871-880(2002).
[2]
RP SEQUENCE OF 210-450 FROM N.A.
RC STRAIN=CD16-1.
RX MEDLINE=21270454; PubMed=11376151;
RA Watanabe T., Miyashita K., Saito T.T., Yoneki T., Kakihara Y.,
RA Nabeshima K., Kishi Y.A., Shimoda C., Nojima H.;
RA "Comprehensive isolation of meiosis-specific genes identifies novel
RA proteins and unusual non-coding transcripts in Schizosaccharomyces
RA pombe.";
RL Nucleic Acids Res. 29:2327-2337(2001).
CC -1- CATALYTIC ACTIVITY: Hydrolysis of terminal 1,4-linked alpha-D-
CC glucose residues successively from non-reducing ends of the chains
CC with release of beta-D-glucose.
CC -1- SIMILARITY: Belongs to family 15 of glycosyl hydrolases.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sb-sib.ch).
CC -----
CC EMBL, AL022305; CAA18423.1; -.
CC EMBL, AB054301; BAB60870.1; -.
CC PIR, T39433; T39433.
CC HSP, P04064; IGAI.
CC GeneDB, SPombe; SPBC14C8.05C; -.
CC InterPro; IPR00165; Glyco_hydro_15.
CC InterPro; IPR008928; Glyco_trans_6hp.
CC Pfam; PF00723; Glyco_hydro_15; 1.

```

RESULT 7
 AMYG_SCHPO
 ID AMYG_SCHPO STANDARD; PRT; 450 AA.
 AC 060087; 096WSS;

DR PRINTS; PR00736; GLHYDLASE15.
 KW Hydrolase; Glycosidase; Polysaccharide degradation; Glycoprotein;
 KM Signal; Meiosis.
 FT SIGNAL 1 16 POTENTIAL.
 FT PROPER 17 28 BY SIMILARITY.
 FT CHAIN 29 450 PROBABLE GLUCAMYLASE.
 FT BINDING 147 147 SUBSTRATE (BY SIMILARITY).
 FT ACT_SITE 203 203 CATALYTIC BASE (BY SIMILARITY).
 FT ACT_SITE 206 206 GENERAL ACID CATALYST (BY SIMILARITY).
 FT ACT_SITE 207 207 INTERACT WITH SUBSTRATES (BY SIMILARITY).
 FT CARBOHYD 383 383 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 409 409 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 450 AA; 51163 MM; 31C5B2ABE4785FE CRC64;
 Query Match 28.2%; Score 781.5; DB 1; Length 450;
 Best Local Similarity 37.6%; Pred. No. 1.9e-44;
 Matches 179; Conservative 68; Mismatches 166; Indels 63; Gaps 12;
 QY 7 LALSGLVCTG---LANVSKRATIDSWLSNATYARTALINNGADGAWGADSGIVVA 63
 DB 8 LLLGGVVASBSLLSPNKRSKASMDWDQOKGLAMGMLNNIGDSGHANDINPGCIIA 67
 QY 64 SPSTDNPDYFYTWTRDSGLVLTVDLPRNGDTSLSTIENYISAQAIIVGGINSPGDL 123
 DB 68 SPSTDSPDYFYQWRDSDALTIMTLDRFEEDKGLPIYKYMBEMWRLQVNPSPGDFY 127
 QY 124 SGAGLSPKFNVDATYATGWSGRQDRGAPALRATAMIGFGQMLDNGYSTATDIWPLV 183
 DB 128 AG-GLGSPKFNVDATYATGWSGRQDRGAPALRATAMIGFGQMLDNGYSTATDIWPLV 186
 QY 184 RNDLSYVAQYVNGQGYDLMEVNGSSPFTTAVQHRALVEGAPATAVGSSCKDSQAP 243
 DB 187 LADLDYTAHMTWTFASFDLMEIKDVHYFTLAVQKRAMODGTAFAKRGIG-----APD 237
 QY 244 -----ILCYLGSFW-TGSEFIANFDS--SRSGKDNATLIGSIHT--PDPEAACD 287
 DB 238 QALAYQRTIEPIDKLGEFMDPGMGVIGYKGRDRSGLDSTLLASISYNEFD----- 291
 QY 288 DSTFQPCSPRALANHKVVDSPFRSIYTLNDGLSDSEAVAGRYEDTYN-----GNPWF 342
 DB 292 -----MHLPYLTLKQETMTDRDYPVNGQMKO-----AMGRYEDVDYDGVSKSIGNPWF 339
 QY 343 LCTLAEEQLYDALYQMDKQSLSEVTVSLDFRK--ALYSMAAGTYSSTSSSTYSYVDA 400
 DB 340 ICSSAAEILYKALAYDKNGLPELTENYHFMKFAEFEGD-----PYNMSVIRKN 390
 QY 401 VKTPADGFVSIVETHAASNGSMSEQYDKSDGEQLSARDLTYSYALLTANNRRNSV 456
 DB 391 MHTYADNFLKAVAFQHPHNGSMSEQFSRDSGHQKARDLTYSYSLNATYRRAI 446
 RESULT 8
 AMYG RHIOR STANDARD; PRT: 604 AA.
 AC P07683;
 DT 01-APR-1988 (Rel. 07, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Glucosylase precursor (EC 3.2.1.3) (Glucan 1,4-alpha-glucosidase)
 OS Rhizopus oryzae (Rhizopus dellemat).
 OC Eukaryota; Fungi; Zygomycota; Zygomycetes; Mucorales; Mucoraceae;
 OC Rhizopus.
 OX NCBI_TaxID=64495;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SAM0034;
 RA Ashikari T., Nakamura N., Tanaka Y., Kiuchi N., Shibano Y.,
 RA Tanaka T., Amachi T., Yoshizumi H.,
 RT "Rhizopus raw-starch-degrading glucosylase: its cloning and
 RT expression in yeast."
 RL Agric. Biol. Chem. 50:957-964(1986).
 RN [2]

RP HOMOLOGY, AND PREDICTED SECONDARY STRUCTURE.
 RA Tanaka Y., Ashikari T., Nakamura N., Kiuchi N., Shibano Y.,
 RA Amachi T., Yoshizumi H.,
 RT "Comparison of amino acid sequences of three glucosylases and their
 RT structure-function relationships."
 RL Agric. Biol. Chem. 50:965-969(1986).
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of terminal 1,4-linked alpha-D-
 CC glucose residues successively from non-reducing ends of the chains
 CC with release of beta-D-glucose.
 CC -1- MISCELLANEOUS: Rhizopus glucosylase exists in multiple forms,
 CC Gluc 1, Gluc 2, and Gluc 3, all of which hydrolyze gelatinized
 CC starch at similar rates, but only the largest one (Gluc 1) is able
 CC to adsorb raw starch.
 CC -1- SIMILARITY: Belongs to family 15 of glycosyl hydrolases.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; D00049; BA00033.1; -.
 CC PIR; JP0001; JP0001.
 DR HSSP; P04064; IGAT.
 DR InterPro; IPR005036; CBM 21.
 DR InterPro; IPR000165; Glyco_hydro_15.
 DR InterPro; IPR008928; Glyco_trans_6hp.
 DR Pfam; PF03370; CBM_21; 1.
 DR Pfam; PF00723; Glyco_hydro_15; 1.
 DR PRINTS; PR00736; GLHYDLASE15.
 DR PROSITE; PS00820; GLUCOMYLASE; 1.
 KW Hydrolase; Glycosidase; Polysaccharide degradation; Glycoprotein;
 KM Signal.
 FT SIGNAL 1 25
 FT CHAIN 26 604 GLUCOMYLASE 1 (GLUC 1).
 FT CHAIN 159 604 GLUCOMYLASE 2 (GLUC 2).
 FT CHAIN 116 604 GLUCOMYLASE 3 (GLUC 3) (OR 110-604).
 FT DOMAIN 26 109 ADSORPTION TO RAW STARCH (OR 26-115).
 FT DOMAIN 116 604 STARCH DEGRADATION (OR 110-604).
 FT BINDING 279 279 SUBSTRATE (BY SIMILARITY).
 FT ACT_SITE 336 336 CATALYTIC BASE (BY SIMILARITY).
 FT ACT_SITE 339 339 GENERAL ACID CATALYST (BY SIMILARITY).
 FT ACT_SITE 340 340 INTERACT WITH SUBSTRATES (BY SIMILARITY).
 SQ SEQUENCE 604 AA; 65162 MM; 78421FLAA3ADB9 CRC64;
 Query Match 26.2%; Score 725; DB 1; Length 604;
 Best Local Similarity 36.8%; Pred. No. 1.6e-40;
 Matches 165; Conservative 71; Mismatches 170; Indels 42; Gaps 10;
 QY 25 ATLDSWLSNEATYARTALINNGADGAWGADSGIVVASPTDNPDPFYTWTRDSGLVL 84
 DB 168 STISSWIKKQEGISRFAMLNINP-----PGSATGFRAALSLTAGPPYYATMRDALTS 222
 QY 85 KTVLDLFR---NGDTSLSTIENYISAQAIIVGGINSPGDLSSGAGLGEKFNVDATY 141
 DB 223 NVIYEVNTTLISGKTKITLANLKDYVTFVSVTQSTSTVCN-----CLGEKFNVDAGYT 276
 QY 142 GSWKRPDRDPALRATAMIGFG--QMLDNGYSTATDIWPLVRNDLSYVAQYVNGQGYD 200
 DB 277 GAWKRPQNDPAPAEATYFLLFADSYLTQYDASVYTGTLKPAFKLDYVNVWVNSGCFD 336
 QY 201 LMEVNGSSPFTTAVQHRALVEGAPATAVGSS--CSMCSQAPLICYQSFWTGSFIL 258
 DB 337 LMEVNGVHYTTLAWMKKGLIGADPAKRKGDSTRASTYSTASTANKISSFFVSSNMW 396
 QY 259 ANFDS-----SRSGKDNATL--LGSIHFPPEAACDSTQPCSPRALANHKVVD 308
 DB 397 IYVSQSVTGAVSKKGLDVLSTLANAGSV-----DDGFFPGSEKILATFAVEDS 447
 QY 309 FRSIYTLNDGLSSEAVAVGRYPEDTY-----YNGNPFICTLAAAEQLYDALYQMDKQ 363


```

Db 448 FASTVPINKNLPVLCNSIGRYPEPTNGNSQSGNFWFLAVTGAELYRAIKEMWING 507
Qy 364 SLEAVDVSLEDFKALYSDAATG-TYSSSSSTYSIVDAKTVADGVSVIETHAASNGM 422
Db 508 GTVSSISLPPFKKDDSSATSGSKTYVTGSDNNIAQNLALADRLSTVQJHAHNGSL 567
Qy 423 SEQYKSDQEQSLARDLTWVSAAALTAN 450
Db 568 AEFPRITGSLTGARDLTWVSAAALTAN 595

RESULT 9
AMYG_SACFI STANDARD; PRT; 519 AA.
ID _AMYG_SACFI STANDARD; PRT; 519 AA.
AC P26989: P78745;
DT 01-AUG-1992 (Rel. 23, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Glucoamylase GLA1 precursor (EC 3.2.1.3) (Glucan 1,4-alpha-
glucosidase) (1,4-alpha-D-glucan glucohydrolase).
CN GLA1.
OS Saccharomycopsis fibuligera (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycopsidaceae; Saccharomycopsis.
OX NCBI_TaxId=4944;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KZ;
RX MEDLINE=92137640; PubMed=1840532;
RA Hostinova E., Balanova J., Gasperik J.;
RT "The nucleotide sequence of the glucoamylase gene GLA1 from
RT Saccharomycopsis fibuligera KZ.";
RL FEMS Microbiol. Lett. 67:103-108(1991).
RN [2]
RP REVISIONS.
RC STRAIN=KZ;
RA Hostinova E.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: Hydrolysis of terminal 1,4-linked alpha-D-
CC glucose residues successively from non-reducing ends of the chains
CC with release of beta-D-glucose.
CC -1- SIMILARITY: BELONGS TO FAMILY 15 OF GLYCOSYL HYDROLASES. ONLY
CC DIFFERS IN 11 POSITIONS WITH GLUCOAMYLASE GLU1.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; X58117; CAA41120.1; -.
CC HSSP; P08017; IAYX.
CC InterPro: IPR000165; Glyco_hydro_15.
CC InterPro: IPR008928; Glyco_trans_1hp.
CC Pfam; PF00723; Glyco_hydro_15; 1.
CC PRINTS; PR00736; GLHYDRLASE15.
CC PROSITE; PS00820; GLUCOAMYLASE, 1.
KW Hydrolase; Glucosidase; Polysaccharide degradation; Glycoprotein;
KW Signal.
FT CHAIN 1 27
FT BINDING 166 519
FT ACT_SITE 234 234 SUBSTRATE (BY SIMILARITY).
FT ACT_SITE 237 237 CATALYTIC BASE (BY SIMILARITY).
FT ACT_SITE 238 238 GENERAL ACID CATALYST (BY SIMILARITY).
FT CARBOHD 115 115 INTERACT WITH SUBSTRATES (BY SIMILARITY).
FT CARBOHD 127 127 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CAROHD 205 205 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 519 AA; 57542 MW; 1A5A009A7640053C CRC64;
Query Match 24.2%; Score 671.5; DB 1; Length 519;

```

```

Best Local Similarity 35.2%; Pred. No. 4,3e-37;
Matches 173; Conservative 73; Mismatches 169; Indels 77; Gaps 16;
Qy 24 RATIDSWLSNEATVARTALINNIGADGAVSGADSGIVASSTNDPDIYFYTWRDSCG- 82
Db 42 RTDLETFLDKQDVLSYVLQNIAYPECGFNDGVGTVIASPSNPDIYYQMTDRDAIT 101
Qy 83 ---VUKTVLDLPRNDTSLSTIENYIGAQAIVQGISNPSG--DLSAGAGLPEPNVDE 137
Db 102 FLTVLSELD--NNFNTLAKAVEYIINTSYNLQRTSNPSGSDDENHGLGEPKFNITG 159
Qy 138 TAYTSGMRPQRDGPALRATAM-----IGFCQMLL-----DNGYTSATDIWPLVR 184
Db 160 SAYTGAMRPPQNDGALALAYALSRILNDVNSLNKGLVITDGDINFST-EDIYKNIIL 218
Qy 185 NDLSTVAOYWNQGYDLMEEVNGSSFFTTAVQRLV-----EGSAPATAVGSSCS 235
Db 219 PLLEVYIGYWDSTGDLWEENGRRHFTSLVQOKLAAVADYAKSFDDGFANTLSSTAS 278
Qy 236 WCDGAPRLCTLOSFTWGS-----FLANFD-----SRSGKANTLIGSIHTFD 281
Db 279 -----TLSEYLSGSDGFPVNTDVNHLVENPDLQONSQGLDSATYIGPLTHD 327
Qy 282 PEACDDSTPQCPSPRALANHKVVDSPRSIYTLNDGLSDSPAVALGRPEPTY-----Y 336
Db 328 -IGESSSTPFVDNRYVLTQSYLLIEDNKRYSVASA--$AGAAIGRPEVYNGDGS 384
Qy 337 NGNPWFCTLAABEOLYQMDKQSLFV--DVSLDFKALYSDAAT----- 384
Db 385 EBNPWFFLATVAAQVPYKLVVD--ASASNDITINKINVDFFKIVYDLSITNGYSSDS 443
Qy 385 GYSSSSSTYSIVDAVTFADGFVSYVETHAASNGSEQYDKSDGELSLARDLTWSA 444
Db 444 VTIKSGDEFTVAVANLVTFGDSFLQVLDHINDGSINBQLNRMTGYSTASYSLTWSG 503
Qy 445 ALLTNRRNSY 456
Db 504 ALLTAIRLRNKY 515

RESULT 10
AMYG_SACFI STANDARD; PRT; 519 AA.
ID _AMYG_SACFI STANDARD; PRT; 519 AA.
AC P08017;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glucoamylase Glu1 precursor (EC 3.2.1.3) (Glucan 1,4-alpha-
glucosidase) (1,4-alpha-D-glucan glucohydrolase).
CN GLU1.
OS Saccharomycopsis fibuligera (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycopsidaceae; Saccharomycopsis.
OX NCBI_TaxId=4944;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=87307999; PubMed=3114236;
RA Itoh T., Ohtsuki I., Yamashita I., Fukui S.;
RT "Nucleotide sequence of the glucoamylase gene Glu1 in the yeast
RT Saccharomycopsis fibuligera.";
RL J. Bacteriol. 169:4171-4176(1987).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS).
RX MEDLINE=98437615; PubMed=9757101;
RA Sevcik J., Solovicova A., Hostinova E., Gasperik J., Wilson K.S.,
RA Dauter Z.;
RT "Structure of glucoamylase from Saccharomycopsis fibuligera at 1.7-A
RT resolution.";
RL Acta Crystallogr. D 54:854-866(1998).
CC -1- CATALYTIC ACTIVITY: Hydrolysis of terminal 1,4-linked alpha-D-
CC glucose residues successively from non-reducing ends of the chains
CC with release of beta-D-glucose.
CC -1- SIMILARITY: BELONGS TO FAMILY 15 OF GLYCOSYL HYDROLASES. ONLY

```

DIFFERS IN 11 POSITIONS WITH GLUCOAMYLASE GLA1.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.ebi.ch/announce/>
CC or send an email to license@ebi-sib.ch).

DR EMBL: L25641; AAA83997.1; -.
DR EMBL: M17355; AAA34649.1; -.
DR PDB: 1AYX; 13-MAY-98.
DR InterPro: IPR00165; Glyco_hydro_15.
DR InterPro: IPR008928; Glyco_trans_6hp.
DR Pfam: PF00723; Glyco_hydro_15; 1.
DR PRINTS: PR00736; GLHYDRLASE15.
DR PROSITE: PS00820; GLUCOAMYLASE; 1.
KW Hydrolase; Glycosidase; Polysaccharide degradation; Glycoprotein;
KW Signal; 3d-structure.
FT SIGNAL 1 27
FT CHAIN 28 519
FT BINDING 166 166
FT ACT_SITE 234 234
FT ACT_SITE 237 237
FT ACT_SITE 238 238
FT CARBOHYD 115 115
FT CARBOHYD 127 127
FT CARBOHYD 205 205
FT TURN 30 31
FT TURN 36 37
FT TURN 42 43
FT HELIX 45 61
FT TURN 62 63
FT STRAND 64 65
FT TURN 67 68
FT STRAND 71 71
FT TURN 76 77
FT STRAND 79 79
FT STRAND 85 85
FT STRAND 89 89
FT STRAND 93 94
FT HELIX 95 111
FT TURN 112 113
FT HELIX 116 133
FT TURN 134 134
FT STRAND 137 137
FT TURN 138 139
FT STRAND 140 140
FT TURN 142 143
FT HELIX 145 151
FT STRAND 154 155
FT TURN 157 158
FT STRAND 161 161
FT TURN 171 171
FT HELIX 173 191
FT TURN 193 194
FT TURN 199 200
FT TURN 202 203
FT HELIX 209 215
FT TURN 216 216
FT HELIX 217 227
FT TURN 228 229
FT STRAND 232 233
FT TURN 235 236
FT STRAND 240 241
FT HELIX 244 263
FT TURN 264 265
FT HELIX 267 285
FT TURN 287 289
FT STRAND 292 293
FT TURN 294 297
FT STRAND 298 299

FT HELIX 303 307
FT TURN 308 309
FT TURN 316 316
FT HELIX 317 325
FT TURN 328 329
FT TURN 338 339
FT HELIX 341 357
FT HELIX 359 363
FT TURN 374 375
FT STRAND 378 378
FT STRAND 383 383
FT STRAND 389 389
FT HELIX 390 410
FT TURN 411 411
FT STRAND 414 417
FT HELIX 418 420
FT HELIX 421 427
FT TURN 428 428
FT HELIX 431 433
FT TURN 434 434
FT TURN 436 437
FT HELIX 438 440
FT STRAND 443 446
FT TURN 448 449
FT HELIX 451 474
FT TURN 477 478
FT STRAND 484 485
FT TURN 487 489
FT STRAND 492 493
FT TURN 496 497
FT HELIX 499 519
SQ SEQUENCE 519 AA; 57539 MW; BE73035AD1B77652 CRC64;

Query Match 24.0%; Score 664.5; DB 1; Length 519;
Best Local Similarity 34.8%; Pred. No. 1.3e-36;
Matches 171; Conservative 73; Mismatches 171; Indels 77; Gaps 16;

24 RATLDWLSNEATVARTAIINIGADGAWVSGADSGIVVASFSTNDPDEYFTWTRDGL- 82
42 RTDLETFRLDKQKEVSLVYLLQNTIAYPEGQFNNGVPGTVIASPSTNSPDYIYQWTRDSAIT 101
83 ---VLKTLVDLFRNGDTSLSITENYISQAIVQGISNPSG--DLSGAGLGEKFKVDE 137
102 FLTVLSELED--NNFNVTLLAKAVEYINTSYNLRQTSNPSGSPDEBNHKGIGSEKFKFTDG 159
138 TAYTGSWGRPORDGPALRATAM-----IGFGWLL-----DNGYSTATDIDVPLVR 184
160 SAYTGAWGRPQNDGPALRAVAISRYYLNDVNSLNEGKVLDTDSGDINFSST-EDIIYKNIIT 218
185 NDLSYVAQWNOGTGYDLMEEVNGSSPFTTIAVQHRALV-----EGSAFATAVAGSSCS 235
219 PDLLEYVIGWDSGTGFDLMEVNGSRHFTTSLVQOKALAYAVDIKSPDDGDFATWLSSTAS 278
236 WCDSSQAPELICYLOSFVTGS-----FILANFP-----SSRSGQDANTLGLSITFPD 281
279 -----TLESYSGSDGGGFVNTDVNHIYENPDLLQONSROGSDSATYIGPLITHD 327
282 PBAACDSTFQPCSPRALANHKEVVDSEFRSITYLLINDGLSDSEAVAVGRYPEDPY-----Y 336
328 -IGSSSTPDPVDNEVYLGSIYLLLEDNKKRYSVNSAY--SAGAAIGRYEDDYNGDSSS 384
337 NGNPWFELCTLAABEQYDALYQWDKQGLEVT--DVSIDPF-----KALYSDPAT 384
385 EGNPMFLATAYAAQVPIKLYD-AKASANDITINKINYDFNKIYVLDSTINSAYQSSDS 443
385 GYSSSSSTYSIYDAVKTFADGFSIVETHAASNGMSQYDKSDGEQLSARDLTWSYA 444
444 VTIKSGSDEFNTVADNLVTFGDSFLQVILDHINDGSLNQLNRKRYTGYSTGAYSLTWSG 503
445 ALLTANNRRNSV 456
504 ALLEAIRLRNKV 515

```

RESULT 11
AMYG_ARYAD STANDARD; PRT; 624 AA.
ID AMYG_ARYAD STANDARD; PRT; 624 AA.
AC P42042;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Glucoamylase precursor (EC 3.2.1.3) (Glucan 1,4-alpha-glucosidase)
DE (1,4-alpha-D-glucan glucohydrolase).
GN GAA.
OS Arxula adeninivorans (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Dipodascaceae; mitosporic Dipodascaceae; Arxula.
OX NCBI_Taxid=37620;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LS3;
RA But Mih D., Kunze I., Foerster S., Wartmann T., Horstmann C.,
RA Mantuffel R., Kunze G.;
RA Submitted (DEC-1994) to the EMBL/GenBank/DDBJ databases.
RL -1- CATALYTIC ACTIVITY: Hydrolysis of terminal 1,4-linked alpha-D-
CC glucose residues successively from non-reducing ends of the chains
CC with release of beta-D-glucose.
CC -1- SIMILARITY: Belongs to family 15 of glycosyl hydrolases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z46901; CA86997.1; -
DR HSSP; P04064; 1GAI.
DR InterPro; IPR005036; CBM_21.
DR InterPro; IPR000165; Glyco_hydro_15.
DR InterPro; IPR008928; Glyco_trans_1.
DR Pfam; PF03370; CBM_21; 1.
DR Pfam; PF00723; Glyco_hydro_15; 1.
DR PRINTS; PR00736; GLHYDRLASE15.
DR PROSITE; PS00820; GLUCOAMYLASE; 1.
KW Hydrolyase; Glycosidase; Polysaccharide degradation; Glycoprotein;
KW signal.
FT SIGNAL 1 18
FT CHAIN 19 624
FT ACT_SITE 340 340 GLUCOAMYLASE.
FT ACT_SITE 343 343 CATALYTIC BASE (BY SIMILARITY).
FT ACT_SITE 344 344 GENERAL ACID CATALYST (BY SIMILARITY).
FT ACT_SITE 344 344 INTERACT WITH SUBSTRATES (BY SIMILARITY).
FT CARBOHD 54 54 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHD 70 70 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHD 98 98 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHD 111 111 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHD 168 168 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHD 267 267 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHD 333 333 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHD 460 460 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHD 582 582 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 624 AA; 68980 MW; 74C2CDBA43FF71B4 CRC64;
Query March 23.2%; Score 641.5; DB 1; Length 624;
Best Local Similarity 32.8%; Pred. No. 53e-35;
Matches 154; Conservative 90; Mismatches 172; Indels 53; Gaps 13;

```

```

143 SWGRPQDGPALRATAMIGFGWLDNG-----YSTATDIPVLENDLSYVAQ 192
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 274 SWGRPQDGPALRATAMIGFGWLDNG-----YSTATDIPVLENDLSYVAQ 192
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 193 YWNGTGYDLMEEVNGSSFFITAVGHRALVEGSAFATAVGSS--CSWCDSQADEILCYLOS 250
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 333 NMSSPSPFLMEEBESAHYTRLVGRKALLLGADFPANDGDHLSDKLTKQASKLSDTLPE 392
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 251 FWYGSFILANFD-----SSRSGKANTLLGSIRTFPEADACDSDTFQPCSPRALANKE 304
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 333 FWDSPARQLILVEYGVLRKGYKSYKDIISVLRGMHGY-----ANDVFSYNDQIILATAYQ 447
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 305 VYDSFRSITYLNDGSDSE---AVAVGRYPEDTY-----YNGNMPFLCTLAALCOLYDA 355
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 448 VTSFLVDYKXANTTSDSGKPLGIPVGRYFEDVDYDGTSGQGNWYLTMMAEFLYRS 507
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 356 LYQMDKQGLSEVTDVSLDFPALYSDA---ATGYSSSSSTYSSTVDVAKTPADGFSVIV 412
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 508 VQEFEDAGSIITISDPLPFWKYFASVDHKGAKYKNKNDQSPKTSLSLTGDAFMRA 567
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 413 ETHAASNSMSEQYDKSGEQLSARDLTWSYALITL-----NNRN 454
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 568 KYHTPSSGHMSEEFNRTGEBRGADLTWSYASLSLAAFAHEELRNQXN 616
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
RESULT 12
AMYG_YEAST STANDARD; PRT; 549 AA.
ID AMYG_YEAST STANDARD; PRT; 549 AA.
AC P08019;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Glucoamylase, intracellular sporulation-specific (EC 3.2.1.3) (Glucan
DE 1,4-alpha-glucosidase) (1,4-alpha-D-glucan glucohydrolase).
GN SGA1 OR SGA OR YII099W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_Taxid=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=87194600; Pubmed=3106330;
RA Yamashita I., Nakamura M., Fukui S.;
RT "Gene fusion is a possible mechanism underlying the evolution of
RT STAL.";
RT J. Bacteriol. 169:2142-2149 (1987).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c / AB972;
RX MEDLINE=97313266; Pubmed=9169870;
RA Churcher C.M., Bowman S., Badcock K., Bankier A., Brown D.,
RA Chillingworth T., Connor R., Devlin K., Gentles S., Hamlin N.,
RA Harris D.E., Horenell T., Hunt S., Jagsels K., Jones M., Lye G.,
RA Moutie S., Odell C., Pearson D., Rajandram M.A., Rice P., Rowley N.,
RA Skelton J., Smith V., Walsh S., Whitehead S., Barrett B.G.,
RL "The nucleotide sequence of Saccharomyces cerevisiae chromosome IX.";
RL Nature 387:84-87 (1997).
CC -1- CATALYTIC ACTIVITY: Hydrolysis of terminal 1,4-linked alpha-D-
CC glucose residues successively from non-reducing ends of the chains
CC with release of beta-D-glucose.
CC -1- SIMILARITY: Belongs to family 15 of glycosyl hydrolases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z38125; CA86282.1; -
DR EMBL; M16166; AAA35042.1; -
DR PIR; S48474; S48474.

```


QY 1 MSFSLALSGVCTGLANVSKRATLSDMSNEATVARTALINNIGADGAVSGADSGI 60
 DB 313 VQLRVLVLMNGTVYVD-SNGAMDSSALEMLOKQKVSIERIFENIGSAVYPS-ILGCV 370
 QY 61 VVASPTNDNDYFYTWTRDGLVLTVDLFRNGDTSLLSTIENYISAQAIYQGISNPSG 120
 DB 371 VIASPSQTHPDYFYQWIMRDSALTINSIVS--HSADPA-IEITLQYLVNSFHLQRTNN--- 424
 QY 121 DLSSGAG-----LGEKFNVDENAYGSMGRPQRDPALRATAMIGCGMLDNG--- 170
 DB 425 TLGAGIGTNDTVAGDPMKVDNNAFTPEWGRPONDDGALLSLILKIIDYIKSGTDL 484
 QY 171 -----YTSTATDIWPLVRNDLSYVAQYWNQGYDLMEVNGSSPFTIAYOHRALVEGSA 225
 DB 485 GAKYFQSGTA-DIPDDIYRMDLRFIDHMNSGPDLMEEVNGMHFTLLVQLSAVDRSL 543
 QY 226 FATAVGSSCSWCD---SQAPILCYL---OSFWTGSFIILNF-----DSRSKGDA 270
 DB 544 YFNASERSSPFVEELRQTRDISKFLVDPANGFINGKY---NYIVETPMIADTLRSGLDI 600
 QY 271 NTLLGSIHTPPEAACDDSTQPCSPRALANKEVDSRSITYTLNDGLSDSEAVAGRY 330
 DB 601 STLAAANTVHADPSA-SHLPEIDINDPAVLTLLHMLHNRSTIYPINDSSKNATGIALGRY 659
 QY 331 PEDTY-----YNGNPFICTLAAAEQLYDALYQMDKQSLSEVTVDSLDFPKALYSDAATG 385
 DB 660 PEDYVDGVEGEGNPMVLATCASTTLVQLYRHSIQHDLVVPANNDCSNAPSELYVPS 719
 QY 386 TYSS-----SSSTYSSIVDAVKTFADGFV 409
 DB 720 NLTTLGNDGYLLEFNTPAFNTQIKIFQLADSF 755

RESULT 14
 AMYI_SACDI STANDARD; PRT; 768 AA.
 ID AC P29760;
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Glucoamylase S2 precursor (EC 3.2.1.3) (Glucon 1,4-alpha-glucosidase)
 DE (1,4-alpha-D-glucan glucohydrolase) (GAI1).
 GN STR2 OR DEX1.
 OS Saccharomyces diastaticus (Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=41870;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94288586; PubMed=8017901;
 RA Kim K., Bajzar G., Lee S.Y., Knudsen F., Mattoon J.R.;
 RT "Cloning of a new allelic variant of a Saccharomyces diastaticus
 RT glucoamylase gene and its introduction into industrial yeasts";
 RL Appl. Biochem. Biotechnol. 44:161-185(1994).
 RN (2)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91276266; PubMed=2055484;
 RA Landreche M.G., Pretorius I.S., Sollietti P., Marmur J.;
 RT "Primary structure and regulation of a glucoamylase-encoding gene
 RT (STW2) in Saccharomyces diastaticus";
 RL Gene 100:95-103(1991).
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of terminal 1,4-linked alpha-D-
 CC glucose residues successively from non-reducing ends of the chains
 CC with release of beta-D-glucose.
 CC -1- SIMILARITY: Belongs to family 15 of glycosyl hydrolases.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL; M60650; AAA5107.1; -
 DR EMBL; M90490; AAA20560.1; -
 DR PIR; J00474; J00474.
 DR HSSP; P08017; IAXY.
 DR InterPro; IPR000165; Glyco_hydro_15.
 DR InterPro; IPR008928; Glyco_trans_6hp.
 DR Pfam; PF00723; Glyco_hydro_15; 1.
 DR PRINTS; PR00736; GHYDR1AS15.
 DR PROSITE; PS00820; GLUCOMYLASE; 1.
 KW Hydrolase; Glycosidase; Polysaccharide degradation; glycoprotein;
 KW Signal; Multigene family.
 FT SIGNAL 1 21
 FT CHAIN 22 768
 FT DOMAIN 22 348
 FT DOMAIN 349 692
 FT BINDING 693 768
 FT ACT_SITE 456 456
 FT ACT_SITE 519 519
 FT ACT_SITE 522 522
 FT ACT_SITE 523 523
 FT CARBOHYD 35 35
 FT CARBOHYD 309 309
 FT CARBOHYD 323 323
 FT CARBOHYD 415 415
 FT CARBOHYD 424 424
 FT CARBOHYD 435 435
 FT CARBOHYD 514 514
 FT CARBOHYD 547 547
 FT CARBOHYD 646 646
 FT CARBOHYD 651 651
 FT CARBOHYD 721 721
 FT CARBOHYD 742 742
 FT CONFLICT 164 164
 FT CONFLICT 624 624
 SQ SEQUENCE 768 AA; 82586 MW; 3FAC172C128A0C6F CRC64;
 D -> N (IN REF. 1).

Query Match 18.8%; Score 521; DB 1; Length 768;

Best Local Similarity 30.5%; Pred. No. 6.4e-27;

Matches 139; Conservative 80; Mismatches 177; Indels 60; Gaps 15;

QY 1 MSFSLALSGVCTGLANVSKRATLSDMSNEATVARTALINNIGADGAVSGADSGI 60
 DB 313 VQLRVLVLMNGTVYVD-SNGAMDSSALEMLOKQKVSIERIFENIGSAVYPS-ILGCV 371
 QY 61 VVASPTNDNDYFYTWTRDGLVLTVDLFRNGDTSLLSTIENYISAQAIYQGISNPSG 120
 DB 371 VIASPSQTHPDYFYQWIMRDSALTINSIVS--HSADPA-IEITLQYLVNSFHLQRTNN--- 425
 QY 121 DLSSGAG-----LGEKFNVDENAYGSMGRPQRDPALRATAMIGCGMLDNG--- 170
 DB 425 TLGAGIGTNDTVAGDPMKVDNNAFTPEWGRPONDDGALLSLILKIIDYIKSGTDL 485
 QY 171 -----YTSTATDIWPLVRNDLSYVAQYWNQGYDLMEVNGSSPFTIAYOHRALVEGSA 225
 DB 485 GAKYFQSGTA-DIPDDIYRMDLRFIDHMNSGPDLMEEVNGMHFTLLVQLSAVDRSL 544
 QY 226 FATAVGSSCSWCD---SQAPILCYL---OSFWTGSFIILNF-----DSRSKGDA 270
 DB 544 YFNASERSSPFVEELRQTRDISKFLVDPANGFINGKY---NYIVETPMIADTLRSGLDI 601
 QY 271 NTLLGSIHTPPEAACDDSTQPCSPRALANKEVDSRSITYTLNDGLSDSEAVAGRY 330
 DB 601 STLAAANTVHADPSA-SHLPEIDINDPAVLTLLHMLHNRSTIYPINDSSKNATGIALGRY 660
 QY 331 PEDTY-----YNGNPFICTLAAAEQLYDALYQMDKQSLSEVTVDSLDFPKALYSDAATG 385
 DB 661 PEDYVDGVEGEGNPMVLATCASTTLVQLYRHSIQHDLVVPANNDCSNAPSELYVPS 720
 QY 386 TYSS-----SSSTYSSIVDAVKTFADGFV 409
 DB 721 NLTTLGNDGYLLEFNTPAFNTQIKIFQLADSF 756

```

RESULT 15
ID YG10 METUA STANDARD; PRT: 615 AA.
AC 059005:
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical glycosyl hydrolase MJ1610 (EC 3.2.1.-).
GN MJ1610.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N. A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8680877;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reisch C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Uetrecht J.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RA "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii."
RL Science 273:1058-1073(1996).
CC -!- SIMILARITY: Belongs to family 15 of glycosyl hydrolases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL: U67601; AAB99630.1; -.
CC DR PIR: A64501; A64501.
CC DR TIGR: MJ1610; -.
CC DR InterPro: IPR000165; Glyco_hydro_15.
CC DR InterPro: IPR008928; Glyco_trans_6hp.
CC DR InterPro: IPR006465; Oligosac_amy1.
CC DR Pfam: PF00723; Glyco_hydro_15; 1.
CC DR TIGRFAMs: TIGR01577; oligosac_amy1; 1.
CC DR PROSITE: PS00820; GUCCAMYLASE; 1.
CC KW Hypothetical protein; Hydrolase; Glycosidase; Complete proteome.
CC FT ACT_SITE 406 406 GENERAL ACID CATALYST (BY SIMILARITY).
CC FT ACT_SITE 407 407 INTERACT WITH SUBSTRATES (BY SIMILARITY).
CC SQ SEQUENCE 615 AA; 72007 MW; 2B37EB89F0357B85 CRC64;

Query Match 8.0%; Score 221.5; DB 1; Length 615;
Best Local Similarity 23.1%; Pred. No. 2.9e-07;
Matches 101; Conservative 46; Mismatches 125; Indels 165; Gaps 20;

QY 59 GIVVSPSTDPDVEFYMTTRSGVLTKLVLPFRNGDTSLSTENYISAQIVQGISNP 118
DB 284 GGIILAPSL-HPDRYRWGRD-GSYISIALDLF-----GINIIPDRFFEFMSKIQ----- 331
QY 119 SGDLSSGAGIGEPKENVDETAVTGSW-----GRPQRDGPALRATAM---IGFGQWLL 167
DB 332 -----NAD-----GSLQNYVYVNGKP-----RLTALIQTDQIGSIILAM 364
QY 168 DNGYTSATDIWPLVLRNDLSYVAQYMNQIG-----YDLMEEVNGSSFF 211
DB 365 DVHYRLR-----GDRKFEVERYNWTLERKANYRLVALNFTPCFDLMEERFFGVFAY 414
QY 212 TIIVQHRLALVGSFAFATV-----GSSCSWCDSQAPRILCYLQSFWTGSIILANFD 262
DB 415 TMGATYAGLKCAYISMKAIVNRKRYKQWKGKTIFFLKHEVPRRP-YLR-----D 461

```

```

QY 263 SSRSGKDANTLLGSIHTPDEAAACDDSTFQPCSPRALANKEVVDSEFRSITYTLNDGLSDS 322
DB 462 EERFAKSIINPL-----DKTIDTSIIIGLSYFPNLIID-----VDDERMIXT-----A 501
QY 323 EAV-----AVGRYPEDTYNNGNPMFLCTLAAA---EQLYDALYQMDKQSLQVTPV 370
DB 502 EALEKAFKRVGIGRYPEDYIFGNNPMITITMLSLYRRLKYKLKENDNGA---DI 557
QY 371 SLDFPVALYSDAATGYSSSSSTYSSIVDAVKTFADGFVSIIVETHAASNGSMSQYDKSD 430
DB 558 YLQSKSKLFWWVKYGF-----DGLFPEQIHKEI 586
QY 431 GEQLSARDLTWSYALAI 447
DB 587 GVPMASAPLQWSNAMFL 603

```

Search completed: June 17, 2004, 17:46:29
Job time : 21 secs